PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 : C12N 15/12, C07K 14/51, 16/22 — A1

(11) International Publication Number: **WO 00/29579**

(43) International Publication Date: 25 May 2000 (25.05.00)

(21) International Application Number: PCT/US99/26909

(22) International Filing Date: 12 November 1999 (12.11.99)

(30) Priority Data:
09/191,986 13 November 1998 (13.11.98) US

(71) Applicant: ZYMOGENETICS, INC. [US/US]; 1201 Eastlake Avenue East, Seattle, WA 98102 (US).

(72) Inventors: LOK, Si; 806 NW 52nd Street, Seattle, WA 98107 (US). PRESNELL, Scott, R.; 2902 N. Puget Sound Avenue, Tacoma, WA 98407 (US).

(74) Agent: LUNN, Paul, G.; ZymoGenetics, Inc., 1201 Eastlake Avenue East, Seattle, WA 98102 (US).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: MAMMALIAN CHONDROMODULIN-LIKE PROTEIN

(57) Abstract

The present invention relates to polynucleotide and polypeptide molecules for mammalian chondromodulin-like polypeptide.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

MAMMALIAN CHONDROMODULIN-LIKE PROTEIN

BACKGROUND OF THE INVENTION

The chondromodulin-I (Chm-I) is a factor described from bovine cartilage [Hiraki *et al. BBRC, 175:* 971-977 (1991), Hiraki *et al. J. Biol. Chem. 271*: 22657-22662 (1996). Chm-I was first reported as an 18 kilodalton (kD) glycoprotein purified from bovine nasal cartilage. The complete amino acid sequence of the purified protein was reported by Neame *et al. J. Biol. Chem. 265*: 9628-9633 (1990). Later, a cDNA encoding what is believed to be a polymorphic variant of this protein, differing by two amino acid residues, was isolated from a bovine fetal epiphysial cartilage cDNA library by Hiraki *et al., BBRC, 175*:971-977 (1991). A human ortholog to bovine Chm-I was discovered by Hirai *et al.* European Patent Application Serial No. EP-624645-A. filed May 11, 1994.

Human and bovine Chm-I share a high degree of structural sequence conservation. Bovine Chm-I is produced as a 335 amino acid precursor polypeptide with features common to the class II cell surface proteins. Most notably, the precursor polypeptide has a nonclassical leader sequence that is followed by a hydrophobic putative transmembrane domain (TMD) sequence. Proteolytic cleavage of the precursor polypeptide gives rise to a mature active protein consisting of the carboxyl-terminal 121 amino acid residues. A processing signal (Arg-Glu-Arg-Arg, SEQ ID NO:10) precedes the mature protein sequence.

Mature bovine and human Chm-I have two domains. One domain covers the amino-terminal half of the mature

protein. The amino-domain is rather hydrophilic and contains all the potential glycosylation sites and exhibits less sequence conservation than the carboxy-terminal domain. The carboxyl-terminal domain contains a hydrophobic region and eight cysteine residues. This domain is highly conserved in bovine and human Chm-I. The residues are 98% identical in this region. The carboxyl-terminal domain of mature Chm-I has been shown to be highly proteinase resistant and is believed to be tightly folded, Neame *et al., J. Biol. Chem. 265*: 9628-9633 (1990). Expression of a cDNA encoding the Chm-I precursor protein in COS-1 cells produces a processed Chm-I protein in the culture supernatant that appears to be structurally and functionally identical to native matured protein.

Bovine *Chm-I* transcripts are found only in cartilaginous tissues. High transcript levels are located in chondrocytes, especially abundant in cells in the proliferating cartilage zones, Hiraki *et al., J. Biol. Chem., 272*: 32419-32426 (1997). *Chm-I* expression is localized specifically to the avascular zones of cartilage in the developing bone and is not expressed in calcifying cartilage nor in bony or the surrounding soft tissues. Immunohistochemical study indicates Chm-I protein co-localizes with its transcript.

Chm-I has a number of bone and cartilage related biological activities. Chm-I has been shown to stimulate osteoblast proliferation, Mori *et al. FEBS Lett., 406*: 310-314 (1997). Bovine Chm-I is shown to stimulate DNA synthesis of culture growth-plate chondrocytes in the presence of basic fibroblast growth factor (FGF), Hiraki *et al., BBRC, 175*: 971-977 (1991). It appears to function as an autocrine chondrocyte colony-stimulating factor, which synergizes with FGF-2 to stimulate colony formation of growth plate chondrocytes in agarose cultures, Inoue *et*

al. BBRC, 241: 395-400 (1997). Most recently, Chm-I has been identified as an endothelial cell growth inhibitor, Hiraki *et al. FEBS Lett, 415*:321-324 (1997) and Hiraki *et al., J. Biol. Chem., 272*: 32419-32426 (1997). This finding suggests that Chm-I has a regulatory role in vascular invasion during endochondral formation and the maintenance of avascularization in cartilage.

There is a need to find additional growth factors which promote the growth of cartilage, muscle, heart and the cells of the mesenchymal tissue lineage in general.

DESCRIPTION OF THE INVENTION

The present invention addresses this need by providing novel polypeptides and related compositions and methods. Within one aspect, the present invention provides an isolated polynucleotide encoding a mammalian chondromodulin-like protein termed Zchm1. The human Zchm1 polypeptide with signal sequence is comprised of a sequence of amino acids 317 amino acids long with the initial Met as shown in SEQ ID NO:1 and SEQ ID NO:2. Amino acid residues 34-48 of SEQ ID NO:2 may define a transmembrane domain, also defined by SEQ ID NO:3. An alternative transmembrane domain may be defined by amino acid residues 31-50 of SEQ ID NO:2, also defined by SEQ ID NO:4. The lysine residues at positions 213-214 indicate a likely cleavage site between amino acid residue 214, a lysine, and amino acid residue 215, a glycine. This would produce a mature Zchm-1 polypeptide extending from amino acid residue 215, a glycine through amino acid residue 317 of SEQ ID NO:2, also define by SEQ ID NO:5. In an alternative embodiment, the mature sequence extends from amino acid residue 255, a phenylalanine, through amino acid residue 317 of SEQ ID NO:2, also defined by SEQ ID NO:6. Other soluble forms of Zchm-1 include a sequence

extending from amino acid residue 202 through amino acid residue 311 of SEQ ID NO:2 also defined by SEQ ID NO:11, and the sequence extending from amino acid residue 48 through amino acid residue 311 of SEQ ID NO:2 also defined by SEQ ID NO:12.

Within a second aspect of the invention there is provided an expression vector comprising (a) a transcription promoter; (b) a DNA segment encoding Zchm1 polypeptide, and (c) a transcription terminator, wherein the promoter, DNA segment, and terminator are operably linked.

Within a third aspect of the invention there is provided a cultured eukaryotic cell into which has been introduced an expression vector as disclosed above, wherein said cell expresses a protein polypeptide encoded by the DNA segment.

Within a further aspect of the invention there is provided a chimeric polypeptide consisting essentially of a first portion and a second portion joined by a peptide bond. The first portion of the chimeric polypeptide consists essentially of (a) a Zchm1 polypeptide as shown in SEQ ID NO: 2 (b) allelic variants of SEQ ID NO:2; and (c) protein polypeptides that are at least 99% identical to (a) or (b). The second portion of the chimeric polypeptide consists essentially of another polypeptide such as an affinity tag. Within one embodiment the affinity tag is an immunoglobulin F_C polypeptide. The invention also provides expression vectors encoding the chimeric polypeptides and host cells transfected to produce the chimeric polypeptides.

Within an additional aspect of the invention there is provided an antibody that specifically binds to a Zchm1 polypeptide as disclosed above, and also an anti-idiotypic

antibody which neutralizes the antibody to a Zchm1 polypeptide.

An additional embodiment of the present invention relates to a peptide or polypeptide which has the amino acid sequence of an epitope-bearing portion of a Zchm1 polypeptide having an amino acid sequence described above. Peptides or polypeptides having the amino acid sequence of an epitope-bearing portion of a Zchm1 polypeptide of the present invention include portions of such polypeptides with at least nine, preferably at least 15 and more preferably at least 30 to 50 amino acids, although epitope-bearing polypeptides of any length up to and including the entire amino acid sequence of a polypeptide of the present invention described above are also included in the present invention. Examples of such polypeptides are polypeptides comprised of SEQ ID NOs: 2, 5, 6, 11, 12 and 16-22. Also claimed are any of these polypeptides that are fused to another polypeptide or carrier molecule.

Prior to setting forth the invention in detail, it may be helpful to the understanding thereof to define the following terms:

The term "affinity tag" is used herein to denote a polypeptide segment that can be attached to a second polypeptide to provide for purification or detection of the second polypeptide or provide sites for attachment of the second polypeptide to a substrate. In principal, any peptide or protein for which an antibody or other specific binding agent is available can be used as an affinity tag. Affinity tags include a poly-histidine tract, protein A, Nilsson *et al.*, *EMBO J.* *4*:1075 (1985); Nilsson *et al.*, *Methods Enzymol.* *198*:3 (1991), glutathione S transferase, Smith and Johnson, *Gene* *67*:31 (1988), Glu-Glu affinity tag, Grussenmeyer *et al.*, *Proc. Natl. Acad. Sci. USA*

82:7952-4 (1985), substance P, Flag™ peptide, Hopp *et al.*, *Biotechnology 6*:1204-1210 (1988), streptavidin binding peptide, or other antigenic epitope or binding domain. See, in general, Ford *et al.*, *Protein Expression and Purification 2*: 95-107 (1991). DNAs encoding affinity tags are available from commercial suppliers (*e.g.*, Pharmacia Biotech, Piscataway, NJ).

The term "allelic variant" is used herein to denote any of two or more alternative forms of a gene occupying the same chromosomal locus. Allelic variation arises naturally through mutation, and may result in phenotypic polymorphism within populations. Gene mutations can be silent (no change in the encoded polypeptide) or may encode polypeptides having altered amino acid sequence. The term allelic variant is also used herein to denote a protein encoded by an allelic variant of a gene.

The terms "amino-terminal" and "carboxyl-terminal" are used herein to denote positions within polypeptides. Where the context allows, these terms are used with reference to a particular sequence or portion of a polypeptide to denote proximity or relative position. For example, a certain sequence positioned carboxyl-terminal to a reference sequence within a polypeptide is located proximal to the carboxyl terminus of the reference sequence, but is not necessarily at the carboxyl terminus of the complete polypeptide.

The term "complement/anti-complement pair" denotes non-identical moieties that form a non-covalently associated, stable pair under appropriate conditions. For instance, biotin and avidin (or streptavidin) are prototypical members of a complement/anti-complement pair. Other exemplary complement/anti-complement pairs include

receptor/ligand pairs, antibody/antigen (or hapten or epitope) pairs, sense/antisense polynucleotide pairs, and the like. Where subsequent dissociation of the complement/anti-complement pair is desirable, the complement/anti-complement pair preferably has a binding affinity of $<10^9$ M^{-1}.

The term "complements of a polynucleotide molecule" is a polynucleotide molecule having a complementary base sequence and reverse orientation as compared to a reference sequence. For example, the sequence 5' ATGCACGGG 3' is complementary to 5' CCCGTGCAT 3'.

The term "contig" denotes a polynucleotide that has a contiguous stretch of identical or complementary sequence to another polynucleotide. Contiguous sequences are said to "overlap" a given stretch of polynucleotide sequence either in their entirety or along a partial stretch of the polynucleotide. For example, representative contigs to the polynucleotide sequence 5'-ATGGCTTAGCTT-3' are 5'-TAGCTTgagtct-3' and 3'-gtcgacTACCGA-5'.

The term "degenerate nucleotide sequence" denotes a sequence of nucleotides that includes one or more degenerate codons (as compared to a reference polynucleotide molecule that encodes a polypeptide). Degenerate codons contain different triplets of nucleotides, but encode the same amino acid residue (i.e., GAU and GAC triplets each encode Asp).

The term "expression vector" is used to denote a DNA molecule, linear or circular, that comprises a segment encoding a polypeptide of interest operably linked to additional segments that provide for its transcription. Such additional segments include promoter and terminator sequences, and may also include one or more origins of

replication, one or more selectable markers, an enhancer, a polyadenylation signal, etc. Expression vectors are generally derived from plasmid or viral DNA, or may contain elements of both.

The term "isolated", when applied to a polynucleotide, denotes that the polynucleotide has been removed from its natural genetic milieu and is thus free of other extraneous or unwanted coding sequences, and is in a form suitable for use within genetically engineered protein production systems. Such isolated molecules are those that are separated from their natural environment and include cDNA and genomic clones. Isolated DNA molecules of the present invention are free of other genes with which they are ordinarily associated, but may include naturally occurring 5' and 3' untranslated regions such as promoters and terminators. The identification of associated regions will be evident to one of ordinary skill in the art (see for example, Dynan and Tijan, *Nature* *316*:774-78 (1985).

An "isolated" polypeptide or protein is a polypeptide or protein that is found in a condition other than its native environment, such as apart from blood and animal tissue. In a preferred form, the isolated polypeptide is substantially free of other polypeptides, particularly other polypeptides of animal origin. It is preferred to provide the polypeptides in a highly purified form, i.e. greater than 95% pure, more preferably greater than 99% pure. When used in this context, the term "isolated" does not exclude the presence of the same polypeptide in alternative physical forms, such as dimers or alternatively glycosylated or derivatized forms.

The term "operably linked", when referring to DNA segments, indicates that the segments are arranged so that they function in concert for their intended purposes,

e.g., transcription initiates in the promoter and proceeds through the coding segment to the terminator.

The term "ortholog" denotes a polypeptide or protein obtained from one species that is the functional counterpart of a polypeptide or protein from a different species. Sequence differences among orthologs are the result of speciation.

"Paralogs" are distinct but structurally related proteins made by an organism. Paralogs are believed to arise through gene duplication. For example, a-globin, b-globin, and myoglobin are paralogs of each other.

A "polynucleotide" is a single- or double-stranded polymer of deoxyribonucleotide or ribonucleotide bases read from the 5' to the 3' end. Polynucleotides include RNA and DNA, and may be isolated from natural sources, synthesized *in vitro*, or prepared from a combination of natural and synthetic molecules. Sizes of polynucleotides are expressed as base pairs (abbreviated "bp"), nucleotides ("nt"), or kilobases ("kb"). Where the context allows, the latter two terms may describe polynucleotides that are single-stranded or double-stranded. When the term is applied to double-stranded molecules it is used to denote overall length and will be understood to be equivalent to the term "base pairs". It will be recognized by those skilled in the art that the two strands of a double-stranded polynucleotide may differ slightly in length and that the ends thereof may be staggered as a result of enzymatic cleavage; thus all nucleotides within a double-stranded polynucleotide molecule may not be paired. Such unpaired ends will in general not exceed 20 nt in length.

A "polypeptide" is a polymer of amino acid residues joined by peptide bonds, whether produced naturally or synthetically. Polypeptides of less than about 10 amino acid residues are commonly referred to as "peptides".

The term "promoter" is used herein for its art-recognized meaning to denote a portion of a gene containing DNA sequences that provide for the binding of RNA polymerase and initiation of transcription. Promoter sequences are commonly, but not always, found in the 5' non-coding regions of genes.

A "protein" is a macromolecule comprising one or more polypeptide chains. A protein may also comprise non-peptidic components, such as carbohydrate groups. Carbohydrates and other non-peptidic substituents may be added to a protein by the cell in which the protein is produced, and will vary with the type of cell. Proteins are defined herein in terms of their amino acid backbone structures; substituents such as carbohydrate groups are generally not specified, but may be present nonetheless.

The term "receptor" denotes a cell-associated protein that binds to a bioactive molecule (i.e., a ligand) and mediates the effect of the ligand on the cell. Membrane-bound receptors are characterized by a multi-domain structure comprising an extracellular ligand-binding domain and an intracellular effector domain that is typically involved in signal transduction. Binding of ligand to receptor results in a conformational change in the receptor that causes an interaction between the effector domain and other molecule(s) in the cell. This interaction in turn leads to an alteration in the metabolism of the cell. Metabolic events that are linked to receptor-ligand interactions include gene transcription, phosphorylation, dephosphorylation,

increases in cyclic AMP production, mobilization of cellular calcium, mobilization of membrane lipids, cell adhesion, hydrolysis of inositol lipids and hydrolysis of phospholipids. In general, receptors can be membrane bound, cytosolic or nuclear; monomeric (*e.g.*, thyroid stimulating hormone receptor, beta-adrenergic receptor) or multimeric (*e.g.*, PDGF receptor, growth hormone receptor, IL-3 receptor, GM-CSF receptor, G-CSF receptor, erythropoietin receptor and IL-6 receptor).

The term "secretory signal sequence" denotes a DNA sequence that encodes a polypeptide (a "secretory peptide") that, as a component of a larger polypeptide, directs the larger polypeptide through a secretory pathway of a cell in which it is synthesized. The larger polypeptide is commonly cleaved to remove the secretory peptide during transit through the secretory pathway.

The term "splice variant" is used herein to denote alternative forms of RNA transcribed from a gene. Splice variation arises naturally through use of alternative splicing sites within a transcribed RNA molecule, or less commonly between separately transcribed RNA molecules, and may result in several mRNAs transcribed from the same gene. Splice variants may encode polypeptides having altered amino acid sequence. The term splice variant is also used herein to denote a protein encoded by a splice variant of an mRNA transcribed from a gene.

Molecular weights and lengths of polymers determined by imprecise analytical methods (*e.g.*, gel electrophoresis) will be understood to be approximate values. When such a value is expressed as "about" X or "approximately" X, the stated value of X will be understood to be accurate to ±10%.

The present invention addresses this need by providing novel polypeptides and related compositions and methods. Within one aspect, the present invention provides an isolated polynucleotide encoding a mammalian chondromodulin-like protein termed Zchm1. The human Zchm1 polypeptide with signal sequence is comprised of a sequence of amino acids 317 amino acids long with the initial Met as shown in SEQ ID NO:1 and SEQ ID NO:2. Amino acid residues 34-48 of SEQ ID NO:2 may define a transmembrane domain, also defined by SEQ ID NO:3. An alternative transmembrane domain may be defined by amino acid residues 31-50 of SEQ ID NO:2, also defined by SEQ ID NO:4. The lysine residues at positions 213-214 indicate a likely cleavage site between amino acid residue 214, a lysine, and amino acid residue 215, a glycine. This would produce a mature Zchm-1 polypeptide extending from amino acid residue 215, a glycine through amino acid residue 317 of SEQ ID NO:2, also define by SEQ ID NO:5. In an alternative embodiment, the mature sequence extends from amino acid residue 255, a phenylalanine, through amino acid residue 317 of SEQ ID NO:2, also defined by SEQ ID NO:6.

The absence of a classical leader sequence and the presence of an amino-terminal hydrophobic sequence, the transmembrane domain, are structural features of the class II cell surface proteins. As such, Zchm-1 may be anchored on the cell membrane *via* its transmembrane domain and oriented with its carboxyl-terminus outwards. Bovine Chm-I precursor protein is believed to adopt the same configuration on the cell surface, Hiraki *et al. BBRC, 175:* 971-977 (1991).

Alignment of bovine and human precursor Chm-I with Zchm-1 sequences reveals a major region of similarity residing

within the carboxy-terminal most 63 amino acid residues. This region of Zchm-1 (residues 255-316) of SEQ ID NO:2 corresponds to the carboxyl-terminal domain of the mature and processed form of bovine and human Chm-I and exhibits 65% sequence identity. Seven of the eight cysteine positions in this region of Zchm-1 (Cys265, Cys266, Cys269, Cys276, Cys280, Cys300 and Cys306) are identical in bovine and human Chm-I, with the single remaining cysteine residue (Cys292) off set by four residue positions toward the amino-terminus in Chm-I. It is predicted that the high degree of sequence conservation along with the near perfect conservation of cysteine positions would result in residues 255-317 adopting one of two possible disulfide bond patterns deduced for bovine Chm-I, Neame *et al.*, *J. Biol. Chem.*, *265*: 9628-9633 (1990).

Amino acid sequence conservation between ZchmI and Chm-I drops off rapidly for sequences preceding Zchm-1 residue 255 of SEQ ID NO:2. The analogous region of bovine and human Chm-I corresponds to the N-terminal domain of the mature protein which is also less conserved between human and bovine Chm-I. The decreased sequence conservation in Chm-I suggests that this region of the protein and perhaps that of Zchm-1 may be functionally less important.

The conserved dibasic cleavage site that give rise to mature bovine and human Chm-I is not present in Zchm-1, an alternative dibasic cleavage site may be present at amino acid residues 213-214 of SEQ ID NO:2. Other than this site there are no obvious alternative processing site in Zchm-1 corresponding to the cleavage site of Chm-I. Accordingly, mature Zchm-1 polypeptide may be a membrane bound protein. There are examples of families of growth factors such as tumor necrosis factors, in which family members are either soluble or are membrane dependent on susceptibility to proteolytic processing.

POLYNUCLEOTIDES:

The present invention also provides polynucleotide molecules, including DNA and RNA molecules, that encode the Zchm1 polypeptides disclosed herein. Those skilled in the art will readily recognize that, in view of the degeneracy of the genetic code, considerable sequence variation is possible among these polynucleotide molecules.

Polynucleotides, generally a cDNA sequence, of the present invention encode the described polypeptides herein. A cDNA sequence which encodes a polypeptide of the present invention is comprised of a series of codons, each amino acid residue of the polypeptide being encoded by a codon and each codon being comprised of three nucleotides. The amino acid residues are encoded by their respective codons as follows.

Alanine (Ala) is encoded by GCA, GCC, GCG or GCT;

Cysteine (Cys) is encoded by TGC or TGT;

Aspartic acid (Asp) is encoded by GAC or GAT;

Glutamic acid (Glu) is encoded by GAA or GAG;

Phenylalanine (Phe) is encoded by TTC or TTT;

Glycine (Gly) is encoded by GGA, GGC, GGG or GGT;

Histidine (His) is encoded by CAC or CAT;

Isoleucine (Ile) is encoded by ATA, ATC or ATT;

Lysine (Lys) is encoded by AAA, or AAG;

Leucine (Leu) is encoded by TTA, TTG, CTA, CTC, CTG or CTT;

Methionine (Met) is encoded by ATG;

Asparagine (Asn) is encoded by AAC or AAT;

Proline (Pro) is encoded by CCA, CCC, CCG or CCT;

Glutamine (Gln) is encoded by CAA or CAG;

Arginine (Arg) is encoded by AGA, AGG, CGA, CGC, CGG or CGT;

Serine (Ser) is encoded by AGC, AGT, TCA, TCC, TCG or TCT;

Threonine (Thr) is encoded by ACA, ACC, ACG or ACT;

Valine (Val) is encoded by GTA, GTC, GTG or GTT;

Tryptophan (Trp) is encoded by TGG; and

Tyrosine (Tyr) is encoded by TAC or TAT.

It is to be recognized that according to the present invention, when a polynucleotide is claimed as described herein, it is understood that what is claimed are both the sense strand, the anti-sense strand, and the DNA as double-stranded having both the sense and anti-sense strand annealed together by their respective hydrogen bonds. Also claimed is the messenger RNA (mRNA) which encodes the polypeptides of the president invention, and which mRNA is encoded by the cDNA described herein. Messenger RNA (mRNA) will encode a polypeptide using the same codons as those defined herein, with the exception that each thymine nucleotide (T) is replaced by a uracil nucleotide (U).

One of ordinary skill in the art will also appreciate that different species can exhibit "preferential codon usage." In general, see, Grantham, *et al.*, *Nuc. Acids Res. 8*:1893-1912 (1980); Haas, *et al. Curr. Biol. 6*:315-324 (1996); Wain-Hobson, *et al.*, *Gene 13*:355-364 (1981); Grosjean and Fiers, *Gene 18*:199-209 (1982); Holm, *Nuc. Acids Res. 14*:3075-3087 (1986); Ikemura, *J. Mol. Biol. 158*:573-597 (1982). As used herein, the term "preferential codon usage" or "preferential codons" is a term of art referring to protein translation codons that are most frequently used in cells of a certain species, thus favoring one or a few

representatives of the possible codons encoding each amino acid. For example, the amino acid Threonine (Thr) may be encoded by ACA, ACC, ACG, or ACT, but in mammalian cells ACC is the most commonly used codon; in other species, for example, insect cells, yeast, viruses or bacteria, different Thr codons may be preferential. Preferential codons for a particular species can be introduced into the polynucleotides of the present invention by a variety of methods known in the art. Introduction of preferential codon sequences into recombinant DNA can, for example, enhance production of the protein by making protein translation more efficient within a particular cell type or species. Sequences containing preferential codons can be tested and optimized for expression in various species, and tested for functionality as disclosed herein.

Within preferred embodiments of the invention the isolated polynucleotides will hybridize to similar sized regions of SEQ ID NO:1, or a sequence complementary thereto, under stringent conditions. In general, stringent conditions are selected to be about 5°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. Typical stringent conditions are those in which the salt concentration is up to about 0.03 M at pH 7 and the temperature is at least about 60°C.

As previously noted, the isolated polynucleotides of the present invention include DNA and RNA. Methods for preparing DNA and RNA are well known in the art. In general, RNA is isolated from a tissue or cell that produces large amounts of Zchm1 RNA. Such tissues and cells are identified by Northern blotting, Thomas, *Proc. Natl. Acad. Sci. USA* 77:5201 (1980), and

include bone marrow and muscle. Total RNA can be prepared using guanidine HCl extraction followed by isolation by centrifugation in a CsCl gradient, Chirgwin *et al.*, *Biochemistry 18*:52-94 (1979). Poly (A)$^+$ RNA is prepared from total RNA using the method of Aviv and Leder, *Proc. Natl. Acad. Sci. USA 69*:1408-1412 (1972). Complementary DNA (cDNA) is prepared from poly(A)$^+$ RNA using known methods. In the alternative, genomic DNA can be isolated. Polynucleotides encoding Zchm1 polypeptides are then identified and isolated by, for example, hybridization or PCR.

A full-length clone encoding Zchm1 can be obtained by conventional cloning procedures. Complementary DNA (cDNA) clones are preferred, although for some applications (*e.g.*, expression in transgenic animals) it may be preferable to use a genomic clone, or to modify a cDNA clone to include at least one genomic intron. Methods for preparing cDNA and genomic clones are well known and within the level of ordinary skill in the art, and include the use of the sequence disclosed herein, or parts thereof, for probing or priming a library. Expression libraries can be probed with antibodies to Zchm1, receptor fragments, or other specific binding partners.

The polynucleotides of the present invention can also be synthesized using DNA synthesizers. Currently the method of choice is the phosphoramidite method. If chemically synthesized double stranded DNA is required for an application such as the synthesis of a gene or a gene fragment, then each complementary strand is made separately. The production of short genes (60 to 80 bp) is technically straightforward and can be accomplished by synthesizing the complementary strands and then annealing them. For the production of longer genes (>300 bp),

however, special strategies must be invoked, because the coupling efficiency of each cycle during chemical DNA synthesis is seldom 100%. To overcome this problem, synthetic genes (double-stranded) are assembled in modular form from single-stranded fragments that are from 20 to 100 nucleotides in length. See Glick and Pasternak, *Molecular Biotechnology, Principles & Applications of Recombinant DNA*, (ASM Press, Washington, D.C. 1994); Itakura *et al.*, *Annu. Rev. Biochem. 53*: 323-356 (1984) and Climie *et al.*, *Proc. Natl. Acad. Sci. USA 87*:633-637 (1990).

The present invention further provides counterpart polypeptides and polynucleotides from other species (orthologs). These species include, but are not limited to mammalian, avian, amphibian, reptile, fish, insect and other vertebrate and invertebrate species. Of particular interest are Zchm1 polypeptides from other mammalian species, including murine, porcine, ovine, bovine, canine, feline, equine, and other primate polypeptides. Orthologs of human Zchm1 can be cloned using information and compositions provided by the present invention in combination with conventional cloning techniques. For example, a cDNA can be cloned using mRNA obtained from a tissue or cell type that expresses Zchm1 as disclosed herein. Suitable sources of mRNA can be identified by probing Northern blots with probes designed from the sequences disclosed herein. A library is then prepared from mRNA of a positive tissue or cell line. A Zchm1-encoding cDNA can then be isolated by a variety of methods, such as by probing with a complete or partial human cDNA or with one or more sets of degenerate probes based on the disclosed sequences. A cDNA can also be cloned using the polymerase chain reaction, or PCR (Mullis, U.S. Patent No. 4,683,202), using primers designed from the representative human Zchm1 sequence

disclosed herein. Within an additional method, the cDNA library can be used to transform or transfect host cells, and expression of the cDNA of interest can be detected with an antibody to Zchm1 polypeptide. Similar techniques can also be applied to the isolation of genomic clones.

Those skilled in the art will recognize that the sequence disclosed in SEQ ID NO:1 represents a single allele of human Zchm1 and that allelic variation and alternative splicing are expected to occur. Allelic variants of this sequence can be cloned by probing cDNA or genomic libraries from different individuals according to standard procedures. Allelic variants of the DNA sequence shown in SEQ ID NO:1, including those containing silent mutations and those in which mutations result in amino acid sequence changes, are within the scope of the present invention, as are proteins which are allelic variants of SEQ ID NO:2. cDNAs generated from alternatively spliced mRNAs, which retain the properties of the Zchm1 polypeptide are included within the scope of the present invention, as are polypeptides encoded by such cDNAs and mRNAs. Allelic variants and splice variants of these sequences can be cloned by probing cDNA or genomic libraries from different individuals or tissues according to standard procedures known in the art.

The present invention also provides isolated Zchm1 polypeptides that are substantially homologous to the polypeptides of SEQ ID NO:2 and their orthologs. The term "substantially homologous" is used herein to denote polypeptides having 50%, preferably 60%, more preferably at least 80%, sequence identity to the sequences shown in SEQ ID NO:2 or their orthologs. Such polypeptides will more preferably be at least 90% identical, and most preferably 95% or more identical to SEQ ID NO:2 or its orthologs.) Percent sequence identity is determined by conventional methods. See, for example, Altschul *et al.*,

Bull. Math. Bio. 48: 603-616 (1986) and Henikoff and Henikoff, *Proc. Natl. Acad. Sci. USA 89*:10915-10919 (1992). Briefly, two amino acid sequences are aligned to optimize the alignment scores using a gap opening penalty of 10, a gap extension penalty of 1, and the "BLOSUM 62" scoring matrix of Henikoff and Henikoff (*ibid.*) as shown in Table 1 (amino acids are indicated by the standard one-letter codes). The percent identity is then calculated as:

$$\frac{\text{Total number of identical matches}}{\text{[length of the longer sequence plus the number of gaps introduced into the longer sequence in order to align the two sequences]}} \times 100$$

Table 1

	A	R	N	D	C	Q	E	G	H	I	L	K	M	F	P	S	T	W	Y	V
A	4																			
R	-1	5																		
N	-2	0	6																	
D	-2	-2	1	6																
C	0	-3	-3	-3	9															
Q	-1	1	0	0	-3	5														
E	-1	0	0	2	-4	2	5													
G	0	-2	0	-1	-3	-2	-2	6												
H	-2	0	1	-1	-3	0	0	-2	8											
I	-1	-3	-3	-3	-1	-3	-3	-4	-3	4										
L	-1	-2	-3	-4	-1	-2	-3	-4	-3	2	4									
K	-1	2	0	-1	-3	1	1	-2	-1	-3	-2	5								
M	-1	-1	-2	-3	-1	0	-2	-3	-2	1	2	-1	5							
F	-2	-3	-3	-3	-2	-3	-3	-3	-1	0	0	-3	0	6						
P	-1	-2	-2	-1	-3	-1	-1	-2	-2	-3	-3	-1	-2	-4	7					
S	1	-1	1	0	-1	0	0	0	-1	-2	-2	0	-1	-2	-1	4				
T	0	-1	0	-1	-1	-1	-1	-2	-2	-1	-1	-1	-1	-2	-1	1	5			
W	-3	-3	-4	-4	-2	-2	-3	-2	-2	-3	-2	-3	-1	1	-4	-3	-2	11		
Y	-2	-2	-2	-3	-2	-1	-2	-3	2	-1	-1	-2	-1	3	-3	-2	-2	2	7	
V	0	-3	-3	-3	-1	-2	-2	-3	-3	3	1	-2	1	-1	-2	-2	0	-3	-1	4

Sequence identity of polynucleotide molecules is determined by similar methods using a ratio as disclosed above.

Variant Zchm1 polypeptides or substantially homologous Zchm1 polypeptides are characterized as having one or more amino acid substitutions, deletions or additions. These changes are preferably of a minor nature, that is conservative amino acid substitutions (see Table 2) and other substitutions that do not significantly affect the folding or activity of the polypeptide; small deletions, typically of one to about 30 amino acids; and small amino- or carboxyl-terminal extensions, such as an amino-terminal methionine residue, a small linker peptide of up to about 20-25 residues, or an affinity tag. The present invention thus includes polypeptides of from 20 to 30 amino acid residues that comprise a sequence that is at least 90%, preferably at least 95%, and more preferably 99% or more identical to the corresponding region of SEQ ID NO:4. Polypeptides comprising affinity tags can further comprise a proteolytic cleavage site between the Zchm1 polypeptide and the affinity tag. Preferred such sites include thrombin cleavage sites and factor Xa cleavage sites.

Table 2

Conservative amino acid substitutions

Basic:	arginine
	lysine
	histidine
Acidic:	glutamic acid
	aspartic acid
Polar:	glutamine
	asparagine
Hydrophobic:	leucine
	isoleucine
	valine
Aromatic:	phenylalanine
	tryptophan
	tyrosine
Small:	glycine
	alanine
	serine
	threonine
	methionine

The present invention further provides a variety of other polypeptide fusions [and related multimeric proteins comprising one or more polypeptide fusions]. For example, a Zchm1 polypeptide can be prepared as a fusion to a dimerizing protein as disclosed in U.S. Patents Nos. 5,155,027 and 5,567,584. Preferred dimerizing proteins in this regard include immunoglobulin constant region domains. Immunoglobulin-Zchm1 polypeptide fusions can be expressed in genetically engineered cells [to produce a variety of multimeric Zchm1 analogs]. Auxiliary domains can be fused to Zchm1 polypeptides to target them to specific cells, tissues, or macromolecules (*e.g.*, collagen). For example, a Zchm1 polypeptide or protein

could be targeted to a predetermined cell type by fusing a Zchm1 polypeptide to a ligand that specifically binds to a receptor on the surface of the target cell. In this way, polypeptides and proteins can be targeted for therapeutic or diagnostic purposes. A Zchm1 polypeptide can be fused to two or more moieties, such as an affinity tag for purification and a targeting domain. Polypeptide fusions can also comprise one or more cleavage sites, particularly between domains. See, Tuan *et al.*, *Connective Tissue Research* *34*:1-9 (1996).

The proteins of the present invention can also comprise non-naturally occurring amino acid residues. Non-naturally occurring amino acids include, without limitation, *trans*-3-methylproline, 2,4-methanoproline, *cis*-4-hydroxyproline, *trans*-4-hydroxyproline, *N*-methylglycine, *allo*-threonine, methylthreonine, hydroxyethylcysteine, hydroxyethylhomocysteine, nitroglutamine, homoglutamine, pipecolic acid, thiazolidine carboxylic acid, dehydroproline, 3- and 4-methylproline, 3,3-dimethylproline, *tert*-leucine, norvaline, 2-azaphenylalanine, 3-azaphenylalanine, 4-azaphenylalanine, and 4-fluorophenylalanine. Several methods are known in the art for incorporating non-naturally occurring amino acid residues into proteins. For example, an *in vitro* system can be employed wherein nonsense mutations are suppressed using chemically aminoacylated suppressor tRNAs.

Methods for synthesizing amino acids and aminoacylating tRNA are known in the art. Transcription and translation of plasmids containing nonsense mutations is carried out in a cell-free system comprising an *E. coli* S30 extract and commercially available enzymes and other reagents. Proteins are purified by chromatography. See,

for example, Robertson *et al.*, *J. Am. Chem. Soc. 113*:2722 (1991); Ellman *et al.*, *Methods Enzymol. 202*:301 (1991; Chung *et al.*, *Science 259*:806-809 (1993); and Chung *et al.*, *Proc. Natl. Acad. Sci. USA 90*:10145-1019 (1993). In a second method, translation is carried out in *Xenopus* oocytes by microinjection of mutated mRNA and chemically aminoacylated suppressor tRNAs, Turcatti *et al.*, *J. Biol. Chem. 271*:19991-19998 (1996). Within a third method, *E. coli* cells are cultured in the absence of a natural amino acid that is to be replaced (*e.g.*, phenylalanine) and in the presence of the desired non-naturally occurring amino acid(s) (*e.g.*, 2-azaphenylalanine, 3-azaphenylalanine, 4-azaphenylalanine, or 4-fluorophenylalanine). The non-naturally occurring amino acid is incorporated into the protein in place of its natural counterpart. See, Koide *et al.*, *Biochem. 33*:7470-7476 (1994). Naturally occurring amino acid residues can be converted to non-naturally occurring species by *in vitro* chemical modification. Chemical modification can be combined with site-directed mutagenesis to further expand the range of substitutions, Wynn and Richards, *Protein Sci. 2*:395-403 (1993).

A limited number of non-conservative amino acids, amino acids that are not encoded by the genetic code, non-naturally occurring amino acids, and unnatural amino acids may be substituted for Zchm1 amino acid residues.

Essential amino acids in the polypeptides of the present invention can be identified according to procedures known in the art, such as site-directed mutagenesis or alanine-scanning mutagenesis, Cunningham and Wells, *Science 244*: 1081-1085 (1989); Bass *et al.*, *Proc. Natl. Acad. Sci. USA 88*:4498-502 (1991). In the

latter technique, single alanine mutations are introduced at every residue in the molecule, and the resultant mutant molecules are tested for biological activity as disclosed below to identify amino acid residues that are critical to the activity of the molecule. See also, Hilton *et al.*, *J. Biol. Chem.* *271*:4699-708, 1996. Sites of ligand-receptor interaction can also be determined by physical analysis of structure, as determined by such techniques as nuclear magnetic resonance, crystallography, electron diffraction or photoaffinity labeling, in conjunction with mutation of putative contact site amino acids. See, for example, de Vos *et al.*, *Science* *255*:306-312 (1992); Smith *et al.*, *J. Mol. Biol.* *224*:899-904 (1992); Wlodaver *et al.*, *FEBS Lett.* *309*:59-64 (1992).

Multiple amino acid substitutions can be made and tested using known methods of mutagenesis and screening, such as those disclosed by Reidhaar-Olson and Sauer, *Science* *241*:53-57 (1988) or Bowie and Sauer, *Proc. Natl. Acad. Sci. USA* *86*:2152-2156 (1989). Briefly, these authors disclose methods for simultaneously randomizing two or more positions in a polypeptide, selecting for functional polypeptide, and then sequencing the mutagenized polypeptides to determine the spectrum of allowable substitutions at each position. Other methods that can be used include phage display, *e.g.*, Lowman *et al.*, *Biochem.* *30*:10832-10837 (1991); Ladner *et al.*, U.S. Patent No. 5,223,409; Huse, WIPO Publication WO 92/06204) and region-directed mutagenesis, Derbyshire *et al.*, *Gene* *46*:145 (1986); Ner *et al.*, *DNA* 7:127 (1988).

Variants of the disclosed Zchm1 DNA and polypeptide sequences can be generated through DNA shuffling as disclosed by Stemmer, *Nature* *370*:389-391,

(1994), Stemmer, *Proc. Natl. Acad. Sci. USA* 91:10747-10751 (1994) and WIPO Publication WO 97/20078. Briefly, variant DNAs are generated by *in vitro* homologous recombination by random fragmentation of a parent DNA followed by reassembly using PCR, resulting in randomly introduced point mutations. This technique can be modified by using a family of parent DNAs, such as allelic variants or DNAs from different species, to introduce additional variability into the process. Selection or screening for the desired activity, followed by additional iterations of mutagenesis and assay provides for rapid "evolution" of sequences by selecting for desirable mutations while simultaneously selecting against detrimental changes.

Mutagenesis methods as disclosed herein can be combined with high-throughput, automated screening methods to detect activity of cloned, mutagenized polypeptides in host cells. Mutagenized DNA molecules that encode active polypeptides can be recovered from the host cells and rapidly sequenced using modern equipment. These methods allow the rapid determination of the importance of individual amino acid residues in a polypeptide of interest, and can be applied to polypeptides of unknown structure.

Using the methods discussed herein, one of ordinary skill in the art can identify and/or prepare a variety of polypeptide fragments or variants of SEQ ID NO:2 or that retain the properties of the wild-type Zchm1 protein.

For any Zchm1 polypeptide, including variants and fusion proteins, one of ordinary skill in the art can readily generate a fully degenerate polynucleotide sequence encoding that variant using the information set forth in Tables 1 and 2 above.

PROTEIN PRODUCTION

The Zchm1 polypeptides of the present invention, including full-length polypeptides, biologically active fragments, and fusion polypeptides, can be produced in genetically engineered host cells according to conventional techniques. Suitable host cells are those cell types that can be transformed or transfected with exogenous DNA and grown in culture, and include bacteria, fungal cells, and cultured higher eukaryotic cells. Eukaryotic cells, particularly cultured cells of multicellular organisms, are preferred. Techniques for manipulating cloned DNA molecules and introducing exogenous DNA into a variety of host cells are disclosed by Sambrook *et al.*, Molecular Cloning: *A Laboratory Manual*, 2nd ed., (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989), and Ausubel *et al.*, eds., *Current Protocols in Molecular Biology* (John Wiley and Sons, Inc., NY, 1987).

In general, a DNA sequence encoding a Zchm1 polypeptide is operably linked to other genetic elements required for its expression, generally including a transcription promoter and terminator, within an expression vector. The vector will also commonly contain one or more selectable markers and one or more origins of replication, although those skilled in the art will recognize that within certain systems selectable markers may be provided on separate vectors, and replication of the exogenous DNA may be provided by integration into the host cell genome. Selection of promoters, terminators, selectable markers, vectors and other elements is a matter of routine design within the level of ordinary skill in the art. Many such elements are described in the literature and are available through commercial suppliers.

To direct a Zchm1 polypeptide into the secretory pathway of a host cell, a secretory signal sequence (also known as a leader sequence, prepro sequence or pre sequence) is provided in the expression vector. The secretory signal sequence may be that of Zchm1, or may be derived from another secreted protein (*e.g.*, t-PA) or synthesized *de novo*. The secretory signal sequence is operably linked to the Zchm1 DNA sequence, i.e., the two sequences are joined in the correct reading frame and positioned to direct the newly synthesized polypeptide into the secretory pathway of the host cell. Secretory signal sequences are commonly positioned 5' to the DNA sequence encoding the polypeptide of interest, although certain secretory signal sequences may be positioned elsewhere in the DNA sequence of interest (see, *e.g.*, Welch *et al.*, U.S. Patent No. 5,037,743; Holland *et al.*, U.S. Patent No. 5,143,830).

Alternatively, the secretory signal sequence contained in the polypeptides of the present invention is used to direct other polypeptides into the secretory pathway. The present invention provides for such fusion polypeptides. The secretory signal sequence contained in the fusion polypeptides of the present invention is preferably fused amino-terminally to an additional peptide to direct the additional peptide into the secretory pathway. Such constructs have numerous applications known in the art. For example, these novel secretory signal sequence fusion constructs can direct the secretion of an active component of a normally non-secreted protein, such as a receptor. Such fusions may be used *in vivo* or *in vitro* to direct peptides through the secretory pathway.

Cultured mammalian cells are suitable hosts within the present invention. Methods for introducing exogenous DNA into mammalian host cells include calcium

phosphate-mediated transfection, Wigler *et al.*, *Cell* *14*:725 (1978); Corsaro and Pearson, *Somatic Cell Genetics* 7:603 (1981); Graham and Van der Eb, *Virology* *52*:456 (1973), electroporation, Neumann *et al.*, *EMBO J.* *1*:841-845 (1982), DEAE-dextran mediated transfection (Ausubel *et al.*, *ibid.*, and liposome-mediated transfection, Hawley-Nelson *et al.*, *Focus* *15*:73 (1993); Ciccarone *et al.*, *Focus* *15*:80 (1993), and viral vectors, Miller and Rosman, *BioTechniques* 7:980(1989); Wang and Finer, *Nature Med.* *2*:714 (1996). The production of recombinant polypeptides in cultured mammalian cells is disclosed, for example, by Levinson *et al.*, U.S. Patent No. 4,713,339; Hagen *et al.*, U.S. Patent No. 4,784,950; Palmiter *et al.*, U.S. Patent No. 4,579,821; and Ringold, U.S. Patent No. 4,656,134. Suitable cultured mammalian cells include the COS-1 (ATCC No. CRL 1650), COS-7 (ATCC No. CRL 1651), BHK (ATCC No. CRL 1632), BHK 570 (ATCC No. CRL 10314), 293 (ATCC No. CRL 1573; Graham *et al.*, *J. Gen. Virol.* *36*:59 (1977) and Chinese hamster ovary (*e.g.* CHO-K1; ATCC No. CCL 61) cell lines. Additional suitable cell lines are known in the art and available from public depositories such as the American Type Culture Collection, Rockville, Maryland. In general, strong transcription promoters are preferred, such as promoters from SV-40 or cytomegalovirus. See, *e.g.*, U.S. Patent No. 4,956,288. Other suitable promoters include those from metallothionein genes (U.S. Patent Nos. 4,579,821 and 4,601,978) and the adenovirus major late promoter.

Drug selection is generally used to select for cultured mammalian cells into which foreign DNA has been inserted. Such cells are commonly referred to as "transfectants". Cells that have been cultured in the presence of the selective agent and are able to pass the

gene of interest to their progeny are referred to as "stable transfectants." A preferred selectable marker is a gene encoding resistance to the antibiotic neomycin. Selection is carried out in the presence of a neomycin-type drug, such as G-418 or the like. Selection systems can also be used to increase the expression level of the gene of interest, a process referred to as "amplification." Amplification is carried out by culturing transfectants in the presence of a low level of the selective agent and then increasing the amount of selective agent to select for cells that produce high levels of the products of the introduced genes. A preferred amplifiable selectable marker is dihydrofolate reductase, which confers resistance to methotrexate. Other drug resistance genes (*e.g.* hygromycin resistance, multi-drug resistance, puromycin acetyltransferase) can also be used. Alternative markers that introduce an altered phenotype, such as green fluorescent protein, or cell surface proteins such as CD4, CD8, Class I MHC, placental alkaline phosphatase may be used to sort transfected cells from untransfected cells by such means as FACS sorting or magnetic bead separation technology.

Other higher eukaryotic cells can also be used as hosts, including plant cells, insect cells and avian cells. The use of *Agrobacterium rhizogenes* as a vector for expressing genes in plant cells has been reviewed by Sinkar *et al.*, *J. Biosci. (Bangalore) 11*:47 (1987). Transformation of insect cells and production of foreign polypeptides therein is disclosed by Guarino *et al.*, U.S. Patent No. 5,162,222 and WIPO publication WO 94/06463. Insect cells can be infected with recombinant baculovirus, commonly derived from Autographa californica nuclear polyhedrosis virus (AcNPV). DNA encoding the Zchm1 polypeptide is inserted into the baculoviral genome in place of the AcNPV polyhedrin gene coding sequence by one

of two methods. The first is the traditional method of homologous DNA recombination between wild-type AcNPV and a transfer vector containing the Zchm1 flanked by AcNPV sequences. Suitable insect cells, *e.g.* SF9 cells, are infected with wild-type AcNPV and transfected with a transfer vector comprising a Zchm1 polynucleotide operably linked to an AcNPV polyhedrin gene promoter, terminator, and flanking sequences. See, King, L.A. and Possee, R.D., *The Baculovirus Expression System: A Laboratory Guide*, (Chapman & Hall, London); O'Reilly, D.R. *et al.*, *Baculovirus Expression Vectors: A Laboratory Manual* (Oxford University Press, New York, New York, 1994); and, Richardson, C. D., Ed., *Baculovirus Expression Protocols. Methods in Molecular Biology*, (Humana Press, Totowa, NJ 1995). Natural recombination within an insect cell will result in a recombinant baculovirus which contains Zchm1 driven by the polyhedrin promoter. Recombinant viral stocks are made by methods commonly used in the art.

The second method of making recombinant baculovirus utilizes a transposon-based system described by Luckow, V.A, *et al.*, *J Virol 67*:4566 (1993). This system is sold in the Bac-to-Bac kit (Life Technologies, Rockville, MD). This system utilizes a transfer vector, pFastBac1™ (Life Technologies) containing a Tn7 transposon to move the DNA encoding the Zchm1 polypeptide into a baculovirus genome maintained in *E. coli* as a large plasmid called a "bacmid." The pFastBac1™ transfer vector utilizes the AcNPV polyhedrin promoter to drive the expression of the gene of interest, in this case Zchm1. However, pFastBac1™ can be modified to a considerable degree. The polyhedrin promoter can be removed and substituted with the baculovirus basic protein promoter (also known as Pcor, p6.9 or MP promoter) which is expressed earlier in the baculovirus infection, and has

been shown to be advantageous for expressing secreted proteins. See, Hill-Perkins, M.S. and Possee, R.D., *J Gen Virol* 71:971 (1990); Bonning, B.C. *et al.*, *J Gen Virol* 75:1551 (1994); and, Chazenbalk, G.D., and Rapoport, B., *J Biol Chem* 270:1543 (1995). In such transfer vector constructs, a short or long version of the basic protein promoter can be used. Moreover, transfer vectors can be constructed which replace the native Zchm1 secretory signal sequences with secretory signal sequences derived from insect proteins. For example, a secretory signal sequence from Ecdysteroid Glucosyltransferase (EGT), honey bee Melittin (Invitrogen, Carlsbad, CA), or baculovirus gp67 (PharMingen, San Diego, CA) can be used in constructs to replace the native Zchm1 secretory signal sequence. In addition, transfer vectors can include an in-frame fusion with DNA encoding an epitope tag at the C- or N-terminus of the expressed Zchm1 polypeptide, for example, a Glu-Glu epitope tag, Grussenmeyer, T. *et al.*, *Proc Natl Acad Sci.* 82:7952 (1985). Using a technique known in the art, a transfer vector containing Zchm1 is transformed into *E. coli*, and screened for bacmids which contain an interrupted lacZ gene indicative of recombinant baculovirus. The bacmid DNA containing the recombinant baculovirus genome is isolated, using common techniques, and used to transfect *Spodoptera frugiperda* cells, *e.g.* Sf9 cells. Recombinant virus that expresses Zchm1 is subsequently produced. Recombinant viral stocks are made by methods commonly used the art.

The recombinant virus is used to infect host cells, typically a cell line derived from the fall army worm, *Spodoptera frugiperda*. See, in general, Glick and Pasternak, *Molecular Biotechnology: Principles and Applications of Recombinant DNA* (ASM Press, Washington, D.C., 1994). Another suitable cell line is the High

FiveO™ cell line (Invitrogen) derived from *Trichoplusia ni* (U.S. Patent #5,300,435). Commercially available serum-free media are used to grow and maintain the cells. Suitable media are Sf900 II™ (Life Technologies) or ESF 921™ (Expression Systems) for the Sf9 cells; and Ex-cell0405™ (JRH Biosciences, Lenexa, KS) or Express FiveO™ (Life Technologies) for the *T. ni* cells. The cells are grown up from an inoculation density of approximately 2-5 x 10^5 cells to a density of 1-2 x 10^6 cells at which time a recombinant viral stock is added at a multiplicity of infection (MOI) of 0.1 to 10, more typically near 3. The recombinant virus-infected cells typically produce the recombinant Zchm1 polypeptide at 12-72 hours post-infection and secrete it with varying efficiency into the medium. The culture is usually harvested 48 hours post-infection. Centrifugation is used to separate the cells from the medium (supernatant). The supernatant containing the Zchm1 polypeptide is filtered through micropore filters, usually 0.45 μm pore size. Procedures used are generally described in available laboratory manuals (King, L. A. and Possee, R.D., *ibid.*; O'Reilly, D.R. *et al.*, *ibid.*; Richardson, C. D., *ibid.*). Subsequent purification of the Zchm1 polypeptide from the supernatant can be achieved using methods described herein.

Fungal cells, including yeast cells, can also be used within the present invention. Yeast species of particular interest in this regard include *Saccharomyces cerevisiae*, *Pichia pastoris*, and *Pichia methanolica*. Methods for transforming *S. cerevisiae* cells with exogenous DNA and producing recombinant polypeptides therefrom are disclosed by, for example, Kawasaki, U.S. Patent No. 4,599,311; Kawasaki *et al.*, U.S. Patent No. 4,931,373; Brake, U.S. Patent No. 4,870,008; Welch *et al.*,

U.S. Patent No. 5,037,743; and Murray *et al.*, U.S. Patent No. 4,845,075. Transformed cells are selected by phenotype determined by the selectable marker, commonly drug resistance or the ability to grow in the absence of a particular nutrient (*e.g.*, leucine). A preferred vector system for use in *Saccharomyces cerevisiae* is the *POT1* vector system disclosed by Kawasaki *et al.* (U.S. Patent No. 4,931,373), which allows transformed cells to be selected by growth in glucose-containing media. Suitable promoters and terminators for use in yeast include those from glycolytic enzyme genes (see, *e.g.*, Kawasaki, U.S. Patent No. 4,599,311; Kingsman *et al.*, U.S. Patent No. 4,615,974; and Bitter, U.S. Patent No. 4,977,092) and alcohol dehydrogenase genes. See also U.S. Patents Nos. 4,990,446; 5,063,154; 5,139,936 and 4,661,454. Transformation systems for other yeasts, including *Hansenula polymorpha, Schizosaccharomyces pombe, Kluyveromyces lactis, Kluyveromyces fragilis, Ustilago maydis, Pichia pastoris, Pichia methanolica, Pichia guillermondii* and *Candida maltosa* are known in the art. See, for example, Gleeson *et al.*, *J. Gen. Microbiol.* *132*:3459 (1986) and Cregg, U.S. Patent No. 4,882,279. *Aspergillus* cells may be utilized according to the methods of McKnight *et al.*, U.S. Patent No. 4,935,349. Methods for transforming *Acremonium chrysogenum* are disclosed by Sumino *et al.*, U.S. Patent No. 5,162,228. Methods for transforming *Neurospora* are disclosed by Lambowitz, U.S. Patent No. 4,486,533.

The use of *Pichia methanolica* as host for the production of recombinant proteins is disclosed in WIPO Publications WO 97/17450, WO 97/17451, WO 98/02536, and WO 98/02565. DNA molecules for use in transforming *P.*

methanolica will commonly be prepared as double-stranded, circular plasmids, which are preferably linearized prior to transformation. For polypeptide production in *P. methanolica*, it is preferred that the promoter and terminator in the plasmid be that of a *P. methanolica* gene, such as a *P. methanolica* alcohol utilization gene (*AUG1* or *AUG2*). Other useful promoters include those of the dihydroxyacetone synthase (DHAS), formate dehydrogenase (FMD), and catalase (CAT) genes. To facilitate integration of the DNA into the host chromosome, it is preferred to have the entire expression segment of the plasmid flanked at both ends by host DNA sequences. A preferred selectable marker for use in *Pichia methanolica* is a *P. methanolica ADE2* gene, which encodes phosphoribosyl-5-aminoimidazole carboxylase (AIRC; EC 4.1.1.21), which allows *ade2* host cells to grow in the absence of adenine. For large-scale, industrial processes where it is desirable to minimize the use of methanol, it is preferred to use host cells in which both methanol utilization genes (*AUG1* and *AUG2*) are deleted. For production of secreted proteins, host cells deficient in vacuolar protease genes (*PEP4* and *PRB1*) are preferred. Electroporation is used to facilitate the introduction of a plasmid containing DNA encoding a polypeptide of interest into *P. methanolica* cells. It is preferred to transform *P. methanolica* cells by electroporation using an exponentially decaying, pulsed electric field having a field strength of from 2.5 to 4.5 kV/cm, preferably about 3.75 kV/cm, and a time constant (t) of from 1 to 40 milliseconds, most preferably about 20 milliseconds.

Prokaryotic host cells, including strains of the bacteria *Escherichia coli*, *Bacillus* and other genera are also useful host cells within the present invention.

Techniques for transforming these hosts and expressing foreign DNA sequences cloned therein are well known in the art, see, *e.g.*, Sambrook *et al.*, *ibid.*). When expressing a Zchm1 polypeptide in bacteria such as *E. coli*, the polypeptide may be retained in the cytoplasm, typically as insoluble granules, or may be directed to the periplasmic space by a bacterial secretion sequence. In the former case, the cells are lysed, and the granules are recovered and denatured using, for example, guanidine isothiocyanate or urea. The denatured polypeptide can then be refolded and dimerized by diluting the denaturant, such as by dialysis against a solution of urea and a combination of reduced and oxidized glutathione, followed by dialysis against a buffered saline solution. In the latter case, the polypeptide can be recovered from the periplasmic space in a soluble and functional form by disrupting the cells (by, for example, sonication or osmotic shock) to release the contents of the periplasmic space and recovering the protein, thereby obviating the need for denaturation and refolding.

Transformed or transfected host cells are cultured according to conventional procedures in a culture medium containing nutrients and other components required for the growth of the chosen host cells. A variety of suitable media, including defined media and complex media, are known in the art and generally include a carbon source, a nitrogen source, essential amino acids, vitamins and minerals. Media may also contain such components as growth factors or serum, as required. The growth medium will generally select for cells containing the exogenously added DNA by, for example, drug selection or deficiency in an essential nutrient which is complemented by the selectable marker carried on the expression vector or co-transfected into the host cell. *P. methanolica* cells are cultured in a medium comprising adequate sources of

carbon, nitrogen and trace nutrients at a temperature of about 25°C to 35°C. Liquid cultures are provided with sufficient aeration by conventional means, such as shaking of small flasks or sparging of fermentors. A preferred culture medium for *P. methanolica* is YEPD (2% D-glucose, 2% Bacto™ Peptone (Difco Laboratories, Detroit, MI), 1% Bacto™ yeast extract (Difco Laboratories), 0.004% adenine and 0.006% L-leucine).

Protein Isolation

It is preferred to purify the polypeptides of the present invention to ≥80% purity, more preferably to ≥90% purity, even more preferably ≥95% purity, and particularly preferred is a pharmaceutically pure state, that is greater than 99.9% pure with respect to contaminating macromolecules, particularly other proteins and nucleic acids, and free of infectious and pyrogenic agents. Preferably, a purified polypeptide is substantially free of other polypeptides, particularly other polypeptides of animal origin.

Expressed recombinant Zchm1 polypeptides (or chimeric Zchm1 polypeptides) can be purified using fractionation and/or conventional purification methods and media. Ammonium sulfate precipitation and acid or chaotrope extraction may be used for fractionation of samples. Exemplary purification steps may include hydroxyapatite, size exclusion, FPLC and reverse-phase high performance liquid chromatography. Suitable chromatographic media include derivatized dextrans, agarose, cellulose, polyacrylamide, specialty silicas, and the like. PEI, DEAE, QAE and Q derivatives are preferred. Exemplary chromatographic media include those media derivatized with phenyl, butyl, or octyl groups, such as Phenyl-Sepharose FF (Pharmacia), Toyopearl butyl 650 (Toso

Haas, Montgomeryville, PA), Octyl-Sepharose (Pharmacia) and the like; or polyacrylic resins, such as Amberchrom CG 71 (Toso Haas) and the like. Suitable solid supports include glass beads, silica-based resins, cellulosic resins, agarose beads, cross-linked agarose beads, polystyrene beads, cross-linked polyacrylamide resins and the like that are insoluble under the conditions in which they are to be used. These supports may be modified with reactive groups that allow attachment of proteins by amino groups, carboxyl groups, sulfhydryl groups, hydroxyl groups and/or carbohydrate moieties. Examples of coupling chemistries include cyanogen bromide activation, N-hydroxysuccinimide activation, epoxide activation, sulfhydryl activation, hydrazide activation, and carboxyl and amino derivatives for carbodiimide coupling chemistries. These and other solid media are well known and widely used in the art, and are available from commercial suppliers. Methods for binding receptor polypeptides to support media are well known in the art. Selection of a particular method is a matter of routine design and is determined in part by the properties of the chosen support. See, for example, *Affinity Chromatography: Principles & Methods* (Pharmacia LKB Biotechnology, Uppsala, Sweden, 1988).

The polypeptides of the present invention can be isolated by exploitation of their properties. For example, immobilized metal ion adsorption (IMAC) chromatography can be used to purify histidine-rich proteins, including those comprising polyhistidine tags. Briefly, a gel is first charged with divalent metal ions to form a chelate, Sulkowski, *Trends in Biochem.* 3:1 (1985). Histidine-rich proteins will be adsorbed to this matrix with differing affinities, depending upon the metal ion used, and will be eluted by competitive elution, lowering the pH, or use of strong chelating agents. Other

methods of purification include purification of glycosylated proteins by lectin affinity chromatography and ion exchange chromatography. *Methods in Enzymol.*, Vol. *182*, "Guide to Protein Purification", M. Deutscher, (ed.),page 529-539 (Acad. Press, San Diego, 1990). Within additional embodiments of the invention, a fusion of the polypeptide of interest and an affinity tag (*e.g.*, maltose-binding protein, an immunoglobulin domain) may be constructed to facilitate purification.

Moreover, using methods described in the art, polypeptide fusions, or hybrid Zchm1 proteins, are constructed using regions or domains of the inventive Zchm1, Sambrook *et al.*, *ibid.*, Altschul *et al.*, *ibid.*, Picard, *Cur. Opin. Biology*, 5:511 (1994). These methods allow the determination of the biological importance of larger domains or regions in a polypeptide of interest. Such hybrids may alter reaction kinetics, binding, constrict or expand the substrate specificity, or alter tissue and cellular localization of a polypeptide, and can be applied to polypeptides of unknown structure.

Fusion proteins can be prepared by methods known to those skilled in the art by preparing each component of the fusion protein and chemically conjugating them. Alternatively, a polynucleotide encoding both components of the fusion protein in the proper reading frame can be generated using known techniques and expressed by the methods described herein. For example, part or all of a domain(s) conferring a biological function may be swapped between Zchm1 of the present invention with the functionally equivalent domain(s) from another family member. Such domains include, but are not limited to, the secretory signal sequence, conserved, and significant domains or regions in this family. Such fusion proteins would be expected to have a biological functional profile

that is the same or similar to polypeptides of the present invention or other known family proteins, depending on the fusion constructed. Moreover, such fusion proteins may exhibit other properties as disclosed herein.

Zchm1 polypeptides or fragments thereof may also be prepared through chemical synthesis. Zchm1 polypeptides may be monomers or multimers; glycosylated or non-glycosylated; pegylated or non-pegylated; and may or may not include an initial methionine amino acid residue.

Chemical Synthesis of Polypeptides

Polypeptides, especially polypeptides of the present invention can also be synthesized by exclusive solid phase synthesis, partial solid phase methods, fragment condensation or classical solution synthesis. The polypeptides are preferably prepared by solid phase peptide synthesis, for example as described by Merrifield, *J. Am. Chem. Soc. 85*:2149 (1963).

ASSAYS

The activity of molecules of the present invention can be measured using a variety of assays. Zchm1 can be measured *in vitro* using cultured cells or *in vivo* by administering molecules of the claimed invention to the appropriate animal model. For instance, Zchm1 transfected (or co-transfected) expression host cells may be embedded in an alginate environment and injected (implanted) into recipient animals. Alginate-poly-L-lysine microencapsulation, permselective membrane encapsulation and diffusion chambers have been described as a means to entrap transfected mammalian cells or primary mammalian cells. These types of non-immunogenic "encapsulations" or microenvironments permit the transfer of nutrients into the microenvironment, and also permit the diffusion of proteins and other macromolecules secreted or released by

the captured cells across the environmental barrier to the recipient animal. Most importantly, the capsules or microenvironments mask and shield the foreign, embedded cells from the recipient animal's immune response. Such microenvironments can extend the life of the injected cells from a few hours or days (naked cells) to several weeks (embedded cells).

An alternative *in vivo* approach for assaying proteins of the present invention involves viral delivery systems. Exemplary viruses for this purpose include adenovirus, herpesvirus, vaccinia virus and adeno-associated virus (AAV). Adenovirus, a double-stranded DNA virus, is currently the best studied gene transfer vector for delivery of heterologous nucleic acid (for a review, see T.C. Becker *et al.*, *Meth. Cell Biol.* *43*:161 (1994); and J.T. Douglas and D.T. Curiel, *Science & Medicine* *4*:44 (1997). The adenovirus system offers several advantages: adenovirus can (i) accommodate relatively large DNA inserts; (ii) be grown to high-titer; (iii) infect a broad range of mammalian cell types; and (iv) be used with a large number of available vectors containing different promoters. Also, because adenoviruses are stable in the bloodstream, they can be administered by intravenous injection.

By deleting portions of the adenovirus genome, larger inserts (up to 7 kb) of heterologous DNA can be accommodated. These inserts can be incorporated into the viral DNA by direct ligation or by homologous recombination with a co-transfected plasmid. In an exemplary system, the essential E1 gene has been deleted from the viral vector, and the virus will not replicate unless the E1 gene is provided by the host cell (the human 293 cell line is exemplary). When intravenously administered to intact animals, adenovirus primarily

targets the liver. If the adenoviral delivery system has an E1 gene deletion, the virus cannot replicate in the host cells. However, the host's tissue (*e.g.*, liver) will express and process (and, if a secretory signal sequence is present, secrete) the heterologous protein. Secreted proteins will enter the circulation in the highly vascularized liver, and effects on the infected animal can be determined.

The adenovirus system can also be used for protein production *in vitro*. By culturing adenovirus-infected non-293 cells under conditions where the cells are not rapidly dividing, the cells can produce proteins for extended periods of time. For instance, BHK cells are grown to confluence in cell factories, then exposed to the adenoviral vector encoding the secreted protein of interest. The cells are then grown under serum-free conditions, which allows infected cells to survive for several weeks without significant cell division. Alternatively, adenovirus vector infected 293S cells can be grown in suspension culture at relatively high cell density to produce significant amounts of protein (see Garnier *et al.*, *Cytotechnol.* *15*:145 (1994). With either protocol, an expressed, secreted heterologous protein can be repeatedly isolated from the cell culture supernatant. Within the infected 293S cell production protocol, non-secreted proteins may also be effectively obtained.

Agonists and Antagonists

In view of the tissue distribution observed for Zchm1, agonists (including the natural ligand/ substrate/ cofactor/ etc.) and antagonists have enormous potential in both *in vitro* and *in vivo* applications. For example, Zchm1 and agonist compounds are useful as components of defined cell culture media, and may be used alone or in

combination with other cytokines and hormones to replace serum that is commonly used in cell culture.

Antagonists

Antagonists are also useful as research reagents for characterizing sites of ligand-receptor interaction. Also as a treatment for prostate cancer. Inhibitors of Zchm1 activity (Zchm1 antagonists) include anti-Zchm1 antibodies and soluble Zchm1 receptors, as well as other peptidic and non-peptidic agents (including ribozymes). Zchm1 can also be used to identify inhibitors (antagonists) of its activity. Test compounds are added to the assays disclosed herein to identify compounds that inhibit the activity of Zchm1. In addition to those assays disclosed herein, samples can be tested for inhibition of Zchm1 activity within a variety of assays designed to measure receptor binding or the stimulation/inhibition of Zchm1-dependent cellular responses. For example, Zchm1-responsive cell lines can be transfected with a reporter gene construct that is responsive to a Zchm1-stimulated cellular pathway. Reporter gene constructs of this type are known in the art, and will generally comprise a Zchm1-DNA response element operably linked to a gene encoding a protein which can be assayed, such as luciferase. DNA response elements can include, but are not limited to, cyclic AMP response elements (CRE), hormone response elements (HRE) insulin response element (IRE), Nasrin *et al.*, *Proc. Natl. Acad. Sci. USA* 87:5273 (1990) and serum response elements (SRE) (Shaw *et al. Cell 56*: 563 (1989). Cyclic AMP response elements are reviewed in Roestler *et al.*, *J. Biol. Chem. 263* (19):9063 (1988) and Habener, *Molec. Endocrinol. 4* (8):1087 (1990). Hormone response elements are reviewed in Beato, *Cell 56*:335 (1989). Candidate compounds, solutions, mixtures or extracts are tested for the ability to inhibit the activity of Zchm1 on the target cells as

evidenced by a decrease in Zchm1 stimulation of reporter gene expression. Assays of this type will detect compounds that directly block Zchm1 binding to cell-surface receptors, as well as compounds that block processes in the cellular pathway subsequent to receptor-ligand binding. In the alternative, compounds or other samples can be tested for direct blocking of Zchm1 binding to receptor using Zchm1 tagged with a detectable label (*e.g.*, ^{125}I, biotin, horseradish peroxidase, FITC, and the like). Within assays of this type, the ability of a test sample to inhibit the binding of labeled Zchm1 to the receptor is indicative of inhibitory activity, which can be confirmed through secondary assays. Receptors used within binding assays may be cellular receptors or isolated, immobilized receptors.

A Zchm1 polypeptide can be expressed as a fusion with an immunoglobulin heavy chain constant region, typically an F_C fragment, which contains two constant region domains and lacks the variable region. Methods for preparing such fusions are disclosed in U.S. Patents Nos. 5,155,027 and 5,567,584. Such fusions are typically secreted as multimeric molecules wherein the Fc portions are disulfide bonded to each other and two non-Ig polypeptides are arrayed in closed proximity to each other. Fusions of this type can be used to affinity purify the ligand. For use in assays, the chimeras are bound to a support via the F_C region and used in an ELISA format.

A Zchm1 ligand-binding polypeptide can also be used for purification of ligand. The polypeptide is immobilized on a solid support, such as beads of agarose, cross-linked agarose, glass, cellulosic resins, silica-based resins, polystyrene, cross-linked polyacrylamide, or like materials that are stable under the conditions of

use. Methods for linking polypeptides to solid supports are known in the art, and include amine chemistry, cyanogen bromide activation, N-hydroxysuccinimide activation, epoxide activation, sulfhydryl activation, and hydrazide activation. The resulting medium will generally be configured in the form of a column, and fluids containing ligand are passed through the column one or more times to allow ligand to bind to the receptor polypeptide. The ligand is then eluted using changes in salt concentration, chaotropic agents (guanidine HCl), or pH to disrupt ligand-receptor binding.

An assay system that uses a ligand-binding receptor (or an antibody, one member of a complement/anti-complement pair) or a binding fragment thereof, and a commercially available biosensor instrument (BIAcore, Pharmacia Biosensor, Piscataway, NJ) may be advantageously employed. Such receptor, antibody, member of a complement/anti-complement pair or fragment is immobilized onto the surface of a receptor chip. Use of this instrument is disclosed by Karlsson, *J. Immunol. Methods* *145*:229 (1991) and Cunningham and Wells, *J. Mol. Biol.* *234*:554 (1993). A receptor, antibody, member or fragment is covalently attached, using amine or sulfhydryl chemistry, to dextran fibers that are attached to gold film within the flow cell. A test sample is passed through the cell. If a ligand, epitope, or opposite member of the complement/anti-complement pair is present in the sample, it will bind to the immobilized receptor, antibody or member, respectively, causing a change in the refractive index of the medium, which is detected as a change in surface plasmon resonance of the gold film. This system allows the determination of on- and off-rates, from which binding affinity can be calculated, and assessment of stoichiometry of binding. Ligand-binding receptor polypeptides can also be used within other assay

systems known in the art. Such systems include Scatchard analysis for determination of binding affinity, Scatchard, *Ann. NY Acad. Sci. 51*: 660 (1949) and calorimetric assays, Cunningham *et al.*, *Science 253*:545 (1991); Cunningham *et al.*, *Science 245*:821 (1991).

Zchm1 polypeptides can also be used to prepare antibodies that specifically bind to Zchm1 epitopes, peptides or polypeptides. The Zchm1 polypeptide or a fragment thereof serves as an antigen (immunogen) to inoculate an animal and elicit an immune response. Suitable antigens would be the Zchm1 polypeptides encoded by SEQ ID NOs:2-24. Antibodies generated from this immune response can be isolated and purified as described herein. Methods for preparing and isolating polyclonal and monoclonal antibodies are well known in the art. See, for example, *Current Protocols in Immunology*, Cooligan, *et al.* (eds.), National Institutes of Health, (John Wiley and Sons, Inc., 1995); Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual, Second Edition* (Cold Spring Harbor, NY, 1989); and Hurrell, J. G. R., Ed., *Monoclonal Hybridoma Antibodies: Techniques and Applications* (CRC Press, Inc., Boca Raton, FL, 1982).

As would be evident to one of ordinary skill in the art, polyclonal antibodies can be generated from inoculating a variety of warm-blooded animals such as horses, cows, goats, sheep, dogs, chickens, rabbits, mice, and rats with a Zchm1 polypeptide or a fragment thereof. The immunogenicity of a Zchm1 polypeptide may be increased through the use of an adjuvant, such as alum (aluminum hydroxide) or Freund's complete or incomplete adjuvant. Polypeptides useful for immunization also include fusion polypeptides, such as fusions of Zchm1 or a portion thereof with an immunoglobulin polypeptide or with maltose

binding protein. The polypeptide immunogen may be a full-length molecule or a portion thereof. If the polypeptide portion is "hapten-like", such portion may be advantageously joined or linked to a macromolecular carrier (such as keyhole limpet hemocyanin (KLH), bovine serum albumin (BSA) or tetanus toxoid) for immunization.

As used herein, the term "antibodies" includes polyclonal antibodies, affinity-purified polyclonal antibodies, monoclonal antibodies, and antigen-binding fragments, such as $F(ab')_2$ and Fab proteolytic fragments. Genetically engineered intact antibodies or fragments, such as chimeric antibodies, Fv fragments, single chain antibodies and the like, as well as synthetic antigen-binding peptides and polypeptides, are also included. Non-human antibodies may be humanized by grafting non-human CDRs onto human framework and constant regions, or by incorporating the entire non-human variable domains (optionally "cloaking" them with a human-like surface by replacement of exposed residues, wherein the result is a "veneered" antibody). In some instances, humanized antibodies may retain non-human residues within the human variable region framework domains to enhance proper binding characteristics. Through humanizing antibodies, biological half-life may be increased, and the potential for adverse immune reactions upon administration to humans is reduced.

Alternative techniques for generating or selecting antibodies useful herein include *in vitro* exposure of lymphocytes to Zchm1 protein or peptide, and selection of antibody display libraries in phage or similar vectors (for instance, through use of immobilized or labeled Zchm1 protein or peptide). Genes encoding polypeptides having potential Zchm1 polypeptide binding domains can be obtained by screening random peptide libraries displayed on phage (phage display) or on

bacteria, such as *E. coli*. Nucleotide sequences encoding the polypeptides can be obtained in a number of ways, such as through random mutagenesis and random polynucleotide synthesis. These random peptide display libraries can be used to screen for peptides which interact with a known target which can be a protein or polypeptide, such as a ligand or receptor, a biological or synthetic macromolecule, or organic or inorganic substances. Techniques for creating and screening such random peptide display libraries are known in the art (Ladner *et al.*, US Patent NO. 5,223,409; Ladner *et al.*, US Patent NO. 4,946,778; Ladner *et al.*, US Patent NO. 5,403,484 and Ladner *et al.*, US Patent NO. 5,571,698) and random peptide display libraries and kits for screening such libraries are available commercially, for instance from Clontech (Palo Alto, CA), Invitrogen Inc. (San Diego, CA), New England Biolabs, Inc. (Beverly, MA) and Pharmacia LKB Biotechnology Inc. (Piscataway, NJ). Random peptide display libraries can be screened using the Zchm1 sequences disclosed herein to identify proteins which bind to Zchm1. These "binding proteins" which interact with Zchm1 polypeptides can be used for tagging cells; for isolating homolog polypeptides by affinity purification; they can be directly or indirectly conjugated to drugs, toxins, radionuclides and the like. These binding proteins can also be used in analytical methods such as for screening expression libraries and neutralizing activity. The binding proteins can also be used for diagnostic assays for determining circulating levels of polypeptides; for detecting or quantitating soluble polypeptides as marker of underlying pathology or disease. These binding proteins can also act as Zchm1 "antagonists" to block Zchm1 binding and signal transduction *in vitro* and *in vivo*. These anti-Zchm1 binding proteins would be useful for inhibiting the effects of Zchm1.

Antibodies are determined to be specifically binding if: (1) they exhibit a threshold level of binding activity, and/or 2) they do not significantly cross-react with related polypeptide molecules. First, antibodies herein specifically bind if they bind to a Zchm1 polypeptide, peptide or epitope with a binding affinity (K_a) of 10^6 M^{-1} or greater, preferably 10^7 M^{-1} or greater, more preferably 10^8 M^{-1} or greater, and most preferably 10^9 M^{-1} or greater. The binding affinity of an antibody can be readily determined by one of ordinary skill in the art, for example, by Scatchard analysis. Second, antibodies are determined to specifically bind if they do not significantly cross-react with related polypeptides. Antibodies do not significantly cross-react with related polypeptide molecules, for example, if they detect Zchm1 but not known related polypeptides using a standard Western blot analysis (Ausubel *et al.*, *ibid.*). Examples of known related polypeptides are orthologs, proteins from the same species that are members of a protein family (*e.g.* IL-16), Zchm1 polypeptides, and non-human Zchm1. Moreover, antibodies may be "screened against" known related polypeptides to isolate a population that specifically binds to the inventive polypeptides. For example, antibodies raised to Zchm1 are adsorbed to related polypeptides adhered to insoluble matrix; antibodies specific to Zchm1 will flow through the matrix under the proper buffer conditions. Such screening allows isolation of polyclonal and monoclonal antibodies non-crossreactive to closely related polypeptides, *Antibodies: A Laboratory Manual*, Harlow and Lane (eds.) (Cold Spring Harbor Laboratory Press, 1988); *Current Protocols in Immunology*, Cooligan, *et al.* (eds.), National Institutes of Health (John Wiley and Sons, Inc., 1995). Screening and isolation of specific antibodies is well known in the

art. See, *Fundamental Immunology*, Paul (eds.) (Raven Press, 1993); Getzoff *et al.*, *Adv. in Immunol.* *43*: 1-98 (1988); *Monoclonal Antibodies: Principles and Practice*, Goding, J.W. (eds.), (Academic Press Ltd., 1996); Benjamin *et al.*, *Ann. Rev. Immunol.* *2*: 67-101 (1984).

A variety of assays known to those skilled in the art can be utilized to detect antibodies which specifically bind to Zchm1 proteins or peptides. Exemplary assays are described in detail in *Antibodies: A Laboratory Manual*, Harlow and Lane (Eds.) (Cold Spring Harbor Laboratory Press, 1988). Representative examples of such assays include: concurrent immunoelectrophoresis, radioimmunoassay, radioimmuno-precipitation, enzyme-linked immunosorbent assay (ELISA), dot blot or Western blot assay, inhibition or competition assay, and sandwich assay. In addition, antibodies can be screened for binding to wild-type versus mutant Zchm1 protein or polypeptide.

Antibodies to Zchm1 may be used for tagging cells that express Zchm1; for isolating Zchm1 by affinity purification; for diagnostic assays for determining circulating levels of Zchm1 polypeptides; for detecting or quantitating soluble Zchm1 as marker of underlying pathology or disease; in analytical methods employing FACS; for screening expression libraries; for generating anti-idiotypic antibodies; and as neutralizing antibodies or as antagonists to block Zchm1 *in vitro* and *in vivo*. Suitable direct tags or labels include radionuclides, enzymes, substrates, cofactors, inhibitors, fluorescent markers, chemiluminescent markers, magnetic particles and the like; indirect tags or labels may feature use of biotin-avidin or other complement/anti-complement pairs as intermediates. Antibodies herein may also be directly or indirectly conjugated to drugs, toxins, radionuclides and

the like, and these conjugates used for *in vivo* diagnostic or therapeutic applications. Moreover, antibodies to Zchm1 or fragments thereof may be used *in vitro* to detect denatured Zchm1 or fragments thereof in assays, for example, Western Blots or other assays known in the art.

BIOACTIVE CONJUGATES:

Antibodies or polypeptides herein can also be directly or indirectly conjugated to drugs, toxins, radionuclides and the like, and these conjugates used for *in vivo* diagnostic or therapeutic applications. For instance, polypeptides or antibodies of the present invention can be used to identify or treat tissues or organs that express a corresponding anti-complementary molecule (receptor or antigen, respectively, for instance). More specifically, Zchm1 polypeptides or anti-Zchm1 antibodies, or bioactive fragments or portions thereof, can be coupled to detectable or cytotoxic molecules and delivered to a mammal having cells, tissues or organs that express the anti-complementary molecule.

Suitable detectable molecules may be directly or indirectly attached to the polypeptide or antibody, and include radionuclides, enzymes, substrates, cofactors, inhibitors, fluorescent markers, chemiluminescent markers, magnetic particles and the like. Suitable cytotoxic molecules may be directly or indirectly attached to the polypeptide or antibody, and include bacterial or plant toxins (for instance, diphtheria toxin, *Pseudomonas* exotoxin, ricin, abrin and the like), as well as therapeutic radionuclides, such as iodine-131, rhenium-188 or yttrium-90 (either directly attached to the polypeptide or antibody, or indirectly attached through means of a chelating moiety, for instance). Polypeptides or antibodies may also be conjugated to cytotoxic drugs, such as adriamycin. For indirect attachment of a detectable or

cytotoxic molecule, the detectable or cytotoxic molecule can be conjugated with a member of a complementary/ anticomplementary pair, where the other member is bound to the polypeptide or antibody portion. For these purposes, biotin/streptavidin is an exemplary complementary/ anticomplementary pair.

In another embodiment, polypeptide-toxin fusion proteins or antibody-toxin fusion proteins can be used for targeted cell or tissue inhibition or ablation (for instance, to treat cancer cells or tissues). Alternatively, if the polypeptide has multiple functional domains (i.e., an activation domain or a ligand binding domain, plus a targeting domain), a fusion protein including only the targeting domain may be suitable for directing a detectable molecule, a cytotoxic molecule or a complementary molecule to a cell or tissue type of interest. In instances where the domain only fusion protein includes a complementary molecule, the anti-complementary molecule can be conjugated to a detectable or cytotoxic molecule. Such domain-complementary molecule fusion proteins thus represent a generic targeting vehicle for cell/tissue-specific delivery of generic anti-complementary-detectable/ cytotoxic molecule conjugates.

In another embodiment, Zchm1-cytokine fusion proteins or antibody-cytokine fusion proteins can be used for enhancing *in vivo* killing of target tissues (for example, blood and bone marrow cancers), if the Zchm1 polypeptide or anti-Zchm1 antibody targets the hyperproliferative blood or bone marrow cell. See, generally, Hornick *et al.*, *Blood 89*:4437 (1997). They described fusion proteins enable targeting of a cytokine to a desired site of action, thereby providing an elevated local concentration of cytokine. Suitable Zchm1 polypeptides or anti-Zchm1 antibodies target an

undesirable cell or tissue (*i.e.*, a tumor or a leukemia), and the fused cytokine mediated improved target cell lysis by effector cells. Suitable cytokines for this purpose include interleukin 2 and granulocyte-macrophage colony-stimulating factor (GM-CSF), for instance.

In yet another embodiment, if the Zchm1 polypeptide or anti-Zchm1 antibody targets vascular cells or tissues, such polypeptide or antibody may be conjugated with a radionuclide, and particularly with a beta-emitting radionuclide, to reduce restenosis. Such therapeutic approach poses less danger to clinicians who administer the radioactive therapy. For instance, iridium-192 impregnated ribbons placed into stented vessels of patients until the required radiation dose was delivered showed decreased tissue growth in the vessel and greater luminal diameter than the control group, which received placebo ribbons. Further, revascularisation and stent thrombosis were significantly lower in the treatment group. Similar results are predicted with targeting of a bioactive conjugate containing a radionuclide, as described herein.

The bioactive polypeptide or antibody conjugates described herein can be delivered intravenously, intraarterially or intraductally, or may be introduced locally at the intended site of action.

USES OF POLYNUCLEOTIDE/POLYPEPTIDE:

The polypeptides of the present invention can be used as a growth or differentiation regulator for cells, especially those of the mesenchymal, myogenic, chondrogenic and endothelial cells. Preparations of the Zchm-1 can be placed in areas in need of bone or cartilage repair. Zchm-1 can also be administered to muscles or endothelial cells which are in need of repair. Instead of

the polypeptide the *Zchm1* gene can be administered as described below.

GENE THERAPY:

Polynucleotides encoding Zchm1 polypeptides are useful within gene therapy applications where it is desired to increase or inhibit Zchm1 activity. If a mammal has a mutated or absent Zchm1 gene, the Zchm1 gene can be introduced into the cells of the mammal. In one embodiment, a gene encoding a Zchm1 polypeptide is introduced *in vivo* in a viral vector. Such vectors include an attenuated or defective DNA virus, such as, but not limited to, herpes simplex virus (HSV), papillomavirus, Epstein Barr virus (EBV), adenovirus, adeno-associated virus (AAV), and the like. Defective viruses, which entirely or almost entirely lack viral genes, are preferred. A defective virus is not infective after introduction into a cell. Use of defective viral vectors allows for administration to cells in a specific, localized area, without concern that the vector can infect other cells. Examples of particular vectors include, but are not limited to, a defective herpes simplex virus 1 (HSV1) vector, Kaplitt *et al.*, *Molec. Cell. Neurosci.* *2*:320 (1991); an attenuated adenovirus vector, such as the vector described by Stratford-Perricaudet *et al.*, *J. Clin. Invest.* *90*:626 (1992); and a defective adeno-associated virus vector, Samulski *et al.*, *J. Virol.* *61*:3096 (1987); Samulski *et al.*, *J. Virol.* *63*:3822 (1989).

In another embodiment, a Zchm1 gene can be introduced in a retroviral vector, *e.g.*, as described in Anderson *et al.*, U.S. Patent No. 5,399,346; Mann *et al.* *Cell* *33*:153, 1983; Temin *et al.*, U.S. Patent No. 4,650,764; Temin *et al.*, U.S. Patent No. 4,980,289;

Markowitz *et al.*, *J. Virol.* *62*:1120 (1988); Temin *et al.*, U.S. Patent No. 5,124,263; International Patent Publication No. WO 95/07358, published March 16, 1995 by Dougherty *et al.*; and Kuo *et al.*, *Blood* *82*:845 (1993). Alternatively, the vector can be introduced by lipofection *in vivo* using liposomes. Synthetic cationic lipids can be used to prepare liposomes for *in vivo* transfection of a gene encoding a marker, Felgner *et al.*, *Proc. Natl. Acad. Sci. USA* *84*:7413 (1987); Mackey *et al.*, *Proc. Natl. Acad. Sci. USA* *85*:8027 (1988). The use of lipofection to introduce exogenous genes into specific organs *in vivo* has certain practical advantages. Molecular targeting of liposomes to specific cells represents one area of benefit. More particularly, directing transfection to particular cells represents one area of benefit. For instance, directing transfection to particular cell types would be particularly advantageous in a tissue with cellular heterogeneity, such as the pancreas, liver, kidney, and brain. Lipids may be chemically coupled to other molecules for the purpose of targeting. Targeted peptides (*e.g.*, hormones or neurotransmitters), proteins such as antibodies, or non-peptide molecules can be coupled to liposomes chemically.

It is possible to remove the target cells from the body; to introduce the vector as a naked DNA plasmid; and then to re-implant the transformed cells into the body. Naked DNA vectors for gene therapy can be introduced into the desired host cells by methods known in the art, *e.g.*, transfection, electroporation, microinjection, transduction, cell fusion, DEAE dextran, calcium phosphate precipitation, use of a gene gun or use of a DNA vector transporter. See, *e.g.*, Wu *et al.*, *J.*

Biol. Chem. *267*:963 (1992); Wu *et al.*, *J. Biol. Chem.* *263*:14621-4, 1988.

Antisense methodology can be used to inhibit Zchm1 gene transcription, such as to inhibit cell proliferation in vivo. Polynucleotides that are complementary to a segment of a Zchm1-encoding polynucleotide (*e.g.*, a polynucleotide as set froth in SEQ ID NO:1) are designed to bind to Zchm1-encoding mRNA and to inhibit translation of such mRNA. Such antisense polynucleotides are used to inhibit expression of Zchm1 polypeptide-encoding genes in cell culture or in a subject.

The present invention also provides reagents which will find use in diagnostic applications. For example, the Zchm1 gene, a probe comprising Zchm1 DNA or RNA or a subsequence thereof can be used to determine if the Zchm1 gene is present on chromosome 11p15.4 or if a mutation has occurred. Detectable chromosomal aberrations at the Zchm1 gene locus include, but are not limited to, aneuploidy, gene copy number changes, insertions, deletions, restriction site changes and rearrangements. Such aberrations can be detected using polynucleotides of the present invention by employing molecular genetic techniques, such as restriction fragment length polymorphism (RFLP) analysis, short tandem repeat (STR) analysis employing PCR techniques, and other genetic linkage analysis techniques known in the art (Sambrook *et al.*, *ibid.*; Ausubel et. al., *ibid.*; Marian, *Chest* *108*:255 (1995).

Transgenic mice, engineered to express the Zchm1 gene, and mice that exhibit a complete absence of Zchm1 gene function, referred to as "knockout mice", Snouwaert *et al.*, *Science* *257*:1083 (1992), may also be generated,

Lowell *et al.*, *Nature 366*:740-42 (1993). These mice may be employed to study the Zchm1 gene and the protein encoded thereby in an *in vivo* system.

CHROMOSOMAL LOCALIZATION:

Radiation hybrid mapping is a somatic cell genetic technique developed for constructing high-resolution, contiguous maps of mammalian chromosomes (Cox *et al.*, *Science 250*:245 (1990). Partial or full knowledge of a gene's sequence allows one to design PCR primers suitable for use with chromosomal radiation hybrid mapping panels. Radiation hybrid mapping panels are commercially available which cover the entire human genome, such as the Stanford G3 RH Panel and the GeneBridge 4 RH Panel (Research Genetics, Inc., Huntsville, AL). These panels enable rapid, PCR-based chromosomal localizations and ordering of genes, sequence-tagged sites (STSs), and other nonpolymorphic and polymorphic markers within a region of interest. This includes establishing directly proportional physical distances between newly discovered genes of interest and previously mapped markers. The precise knowledge of a gene's position can be useful for a number of purposes, including: 1) determining if a sequence is part of an existing contig and obtaining additional surrounding genetic sequences in various forms, such as YACs, BACs or cDNA clones; 2) providing a possible candidate gene for an inheritable disease which shows linkage to the same chromosomal region; and 3) cross-referencing model organisms, such as mouse, which may aid in determining what function a particular gene might have. *ZChm-1* was mapped to chromosome 11p15.4 region.

Sequence tagged sites (STSs) can also be used independently for chromosomal localization. An STS is a DNA sequence that is unique in the human genome and can be

used as a reference point for a particular chromosome or region of a chromosome. An STS is defined by a pair of oligonucleotide primers that are used in a polymerase chain reaction to specifically detect this site in the presence of all other genomic sequences. Since STSs are based solely on DNA sequence they can be completely described within an electronic database, for example, Database of Sequence Tagged Sites (dbSTS), GenBank, (National Center for Biological Information, National Institutes of Health, Bethesda, MD http://www.ncbi.nlm.nih.gov), and can be searched with a gene sequence of interest for the mapping data contained within these short genomic landmark STS sequences.

For pharmaceutical use, the proteins of the present invention are formulated for parenteral, particularly intravenous or subcutaneous, delivery according to conventional methods. Intravenous administration will be by bolus injection or infusion over a typical period of one to several hours. In general, pharmaceutical formulations will include a Zchm1 protein in combination with a pharmaceutically acceptable vehicle, such as saline, buffered saline, 5% dextrose in water or the like. Formulations may further include one or more excipients, preservatives, solubilizers, buffering agents, albumin to prevent protein loss on vial surfaces, etc. Methods of formulation are well known in the art and are disclosed, for example, in Remington: *The Science and Practice of Pharmacy*, Gennaro, ed., (Mack Publishing Co., Easton, PA, 19th ed., 1995). Therapeutic doses will generally be in the range of 0.1 to 100 µg/kg of patient weight per day, preferably 0.5-20 mg/kg per day, with the exact dose determined by the clinician according to accepted standards, taking into account the nature and severity of the condition to be treated, patient traits, etc. Determination of dose is within the level of

ordinary skill in the art. The proteins may be administered for acute treatment, over one week or less, often over a period of one to three days or may be used in chronic treatment, over several months or years. In particular these preparations can be used to promote bone or cartilage repair.

RESEARCH TOOL UTILITY

The polynucleotides provided by the present invention can be used by the research community for various purposes. The polynucleotides can be used to express recombinant protein for analysis, characterization or therapeutic; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or disease states); as molecular weight markers on Southern gels; as chromosome markers (when labeled) to map related gene positions; to compare with endogenous DNA sequences in patients to identify potential genetic disorders; as probes to hybridize and thus discover novel, related DNA sequences; as a source of information to derive PCR primers for genetic fingerprinting; as a probe to "subtract-out" known sequences in the process of discovering other novel polynucleotides; to raise anti-protein antibodies using DNA immunization techniques; and as an antigen to raise anti-DNA antibodies or elicit another immune response. Where the polynucleotide encodes a protein which binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the polynucleotide can also be used in interaction trap assays [such as, for example, that described in Gyuris *et al.* *Cell* 75:791-803 (1993)] to identify polynucleotides

encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction.

The proteins provided by the present invention can similarly be used to raise antibodies or to elicit another immune response: as a reagent (including the labeled reagent) in assays designed to quantitatively determine levels of the protein (or its receptor) in biological fluids; as markers for tissues using labeled antibodies; and to isolate correlative receptors or ligands. Where the protein binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the protein can be used to identify the other protein with which binding occurs or to identify inhibitors of the binding interaction. Proteins involved in these binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction.

Any or all of these "research tool" utilities are capable of being developed into reagent grade or kit format for commercialization as "research products".

Cytokine and Cell Proliferation/Differentiation Activity

A protein of the present invention may exhibit cytokine-cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations. Many protein factors discovered to date, including all know cytokines, have exhibited activity in one or more factor dependent cell proliferation assays, and hence the assays serve has a convenient confirmation of cytokine activity. The activity of a protein of the present invention is evidenced by any one of a number of routine factor dependent cell proliferation assays for cell lines

including without limitation, 32D, DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M+(preB M+), 2E8, RB5, DA1,123, T1165, HT2, CTLL2, TF-1, Mo7e and CMK.

The activity of a protein of the invention may, among other means, be measured by assays for T-cell or thymocyte proliferation, assays for cytokine production or proliferation of spleen cells, lymph node cells or thymocytes, assays for proliferation and differentiation of hematopoietic and lymphopoietic cells, and assays for T-cell clone responses to antigens which will identify, among others, proteins that affect antigen-presenting cells (APC)/T-cell interactions as well as direct T-cell effects by measuring proliferation and cytokine production. Other immunological assays include assays for T-cell dependent immunoglobulin responses and isotype switching (which will identify, among others, proteins that modulate T-cell dependent antibody responses and that affect Th1/Th2 profiles); mixed lymphocyte reaction (MLR) assays (which will identify proteins that generate predominantly Th1 and CTL responses); dendritic cell-dependent assays (which will identify, among others, proteins expressed a by dendritic cells that activate naïve T-cells); assays for lymphocyte survival/apoptosis (which will identify proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis);assays for B cell function and assays for protein that influence early steps of T-cell commitment and development. The above-described assays are described in one or more of the following references: *Current Protocols in Immunology*, (John Wiley and Sons, Toronto, 1997); Takai *et al.*, *J. Immunol. 137*:3494-3500 (1986); Bertagnolli *et al. J. Immunol. 145*:1706-1712 (1990); Bertagnolli *et al.*, *Cell. Immunol. 133*:327-341 (1991); Bertagnolli *et al.*, *J. Immunol. 149*:3778-3783 (1992); Bowman *et al.*, *J. Immunol. 152*:1756-1761 (1994);

de Vries *et al.*, *J. Exp. Med.* *173*:1205-1211 (1991); Moreau *et al.*, *Nature* *336*:690-692 (1988); Greenberger *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* *80*:2931-2938 (1983); Weinberger *et al.*, *Proc. Natl. Acad. Sci. USA*, *77*:6091-6095 (1980); Weinberger *et al.*, *Eur. J. Immunol.* *11*:405-411 (1981); Takai *et al.*, *J. Immunol.* *140*:508-512 (1988); Maliszewski, *J. Immunol.* *144:* 3028-3033 (1990); Herrmann *et al.*, *Proc. Natl Acad. Sci USA* *78*:24882492 (1981); Herrmann *et al.*, *J. Immunol.* *128*:1968-1974 (1982); Handa *et al.* *J. Immunol.* *135*:1564-1572 (1985); Bowmanet *et al.*, *J. Virology* *61*:1992-1998; Brown *et al.*, *J. Immunol.* *153*:3079-3092 (1994); Maliszewski, *J. Immunol.* *144*:3028-3033 (1990); Guery *et al.* *J. Immunol.* *134*:536-544 (1995); Inaba *et al.*, *J. Exp. Med.* *173*:549-559 (1991); Macatonia *et al.*, *J. Immunol.* *154*:5071-5079 (1995); Porgador *et al.*, *J. Exp. Med.* *182*:255-260 (1995); Nair *et al.* *J. Virol.* *67*:4062-4069 (1993); Huang *et al.*, *Science* *264*:961-965 (1994); Macatonia *et al.*, *J. Exp. Med.* *169*:1255-1264 (1989); Bhardwaj *et al.*, *J. Clin. Invest.* *94*:797-807 (1994); Inaba *et al.*, *J. Exp. Med.* *172*:631-640 (1990); Darzynkiewicz *et al.*, *Cytometry* *13*:795-808 (1992); Gorczyca *et al.*, *Leukemia* *7*:659-670 (1993); Gorczyca *et al.*, *Can. Res.* *53*:1945-1951 (1993); Itoh *et al.*, *Cell* *66*:233-243 (1991);Zacharchuk, *J. Immunol.* *145*:4037-4045 (1990); Zamai *et al.* *Cytometry* *14*:891-897 (1993); Gorczyca *et al.*, *Inter. J. Oncol.* *1*:639-648 (1992); Antica *et al.*, *Blood* *84*:111-117 (1994); Fine *et al.*, *Cell. Immunol.* *155*:111-122 (1994); Galy *et al.*, *Blood* *85*:2770-2778 (1995); and Toki *et al.*, *Proc. Natl. Acad Sci. USA* *88*:7548-7551 (1991).

Immune Stimulating/Suppressing Activity

A protein of the present invention may also exhibit immune stimulating or immune suppressing activity including, without limitation, the activities for which assays are described herein. A protein may be useful in the treatment of various immune deficiencies and disorders [including severe combined immunodeficiency (SCIC)], *e.g.*, in regulating (up or down) growth and proliferation of T or B lymphocytes., as well as effecting the cytolytic activity of natural killer (NK) cells and other cell populations. These immune deficiencies may be genetic or by caused by viral as well as bacterial or fungal infections or may result from autoimmune disorders. The protein of the present invention by may possibly be used to treat such diseases or to boost the immune system.

Hematopoiesis

The protein of the present invention may be useful in promoting hematopoiesis, including causing proliferation of red blood cells, megakaryocytes, and myeloid cells such as monocytes/macrophages. Assays for relating to stem cell growth or differentiation include: Freshney, M.G., in *Culture of Hematopoietic Cells*, Frshney, R.I. *et al.*, *Eds.* (Wiley-Liss, Inc., New York, N.Y., 1994); Johansson *et al. Cell. Bio. 15*:141-151 (1995); Keller *et al.*, *Mol. & Cell. Bio. 13*:473-486 (1993); McClanahan *et al.*, *Blood 81*:2903-2915 (1993); Hirayama *et al.*, *Proc. Natl. Acad. Sci. USA 89*:5907-5911 (1992); and Neben *et al.*, *Exp. Hematol. 22*:353-359 (1994).

Tissue Regeneration or Repair

The protein of the present invention may be used to repair or regenerate any number of different tissues including bone, ligaments, tendons, neurons and skin. Assays for tissue regeneration include those described in International Patent Publication No. WO95/16035 (bone, cartilage, tendon); WO95/05846 (neuron); and WO91/07491 (skin, endothelium).

Activin/Inhibin Activity

A protein of the present invention may also exhibit activin or inhibin related activities. Inhibin is a glycoprotein that circulates in plasma and inhibits gonadotropin-releasing hormone (GnRH)-stimulated follicle stimulating hormone (FSH) secretion by the pituitary gland. Activin has the opposite action and stimulates FSH secretion. Thus, the protein of the present invention may be useful as a contraceptive or as a based upon the

ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Assays for activin/inhibin activity are described in the following: Vale *et al.*, *Endocrinology 91*:562-572 (1972); Ling *et al.*, *Nature 321:* 779-782 (1986); Vale *et al.*, *Nature 321*:776-779 (1986); Mason *et al.*, *Nature 318*:659-663 (1985); Forage *et al.*, *Proc. Natl. Acad. Sci. USA 83*:3091-3095 (1986).

The invention is further illustrated by the following non-limiting examples.

Example 1

Discovery of Zchm1 and Construction of Expression Plasmid

Zchm1 cDNA was discovered in a mesentery tumor, sigmoid, mets mixed-mullerian tumor cDNA library. The library was constructed using 1 microgram of polyA RNA isolated from sigmoid mesentery tumor tissue removed from a 61-year-old female during a total abdominal hysterectomy and bilateral salpingo-oophorectomy with regional lymph node excision. Pathology indicated a metastatic grade 4 malignant mixed mullerian tumor present in the sigmoid mesentery at two sites. This tumor was associated with a grade 4 malignant mixed-mullerian tumor, heterologous type, of the uterus, forming a firm, infiltrating mass throughout the myometrium and involving the serosal surface. The heterologous elements of the tumor consisted of rhabdomyoblasts and immature cartilage. The tumor also involved the lower uterine segment and extended into the cervical wall. Extensive lymphatic and vascular permeation was identified in the myometrium and cervical wall. One (of 7) right common iliac and one (of 7) right external iliac lymph nodes were identified with metastatic grade 4 malignant mixed mullerian tumor, with the metastases

comprised mainly of adenocarcinoma. There were also positive estrogen and progesterone receptors. cDNA synthesis was initiated using a NotI-oligo(dT) primer. Double-stranded cDNA was blunted, ligated to EcoRI adaptors, digested with NotI, size-selected, and cloned into the NotI and EcoRI sites a plasmid. Expressed sequence tags (EST) from the cDNA library were scanned to discover sequences having homology to the *chondromodulin* gene. The EST of SEQ ID NO:15 was discovered and the clone obtained and sequenced resulting in the *Zchm1* gene of SEQ ID NO:1. The open reading frame of *Zchm-1* was subcloned into vector *pZP9CEE* for expression in mammalian cells. The *pZP9CEE* was the *pZP9* (ATCC 98668) into which DNA encoding a glu-glu epitope tag was inserted. See Grussenmeyer, T. *et al.*, *Proc Natl Acad Sci. 82*:7952 (1985). To facilitate cloning into the vector, a polymerase chain reaction was used to create a *Sal* I and a *Bam* HI sites upstream from the *Zchm-1* initiation codon and downstream from the translation termination codon, respectively. The polymerase chain reaction was carried out in a reaction containing 10 µl of native *Pfu* DNA polymerase buffer, 1 µl of a deoxynucleotide triphosphate solution containing 10 mM each of dATP, dGTP, dTTP and dCTP (Pharmacia LKB Biotechnology Inc., Piscataway, NJ), 5 µl of 7 pmole/µl primer ZC18149 (SEQ ID NO:7), 5 µl of 7 pmole/µl primer ZC18150 (SEQ ID NO: 8), 74 µl of water, 3 µl of 10 ng/µl *Xho* I digested *pSLZchm1-3* template, a plasmid which contained the *Zchm-1* cDNA, and 2 µl of 2.5 U/µl native *Pfu* DNA polymerase (Stratagene Cloning Systems, La Jolla, CA). The polymerase chain reaction was run for 13 cycles (20 seconds at 95° C and 1.5 minutes at 72° C) followed by a 5 minute incubation at 72° C. The amplified products were serially extracted with

phenol/chloroform, chloroform and precipitated with ethanol in the presence of PELLET PAINT ® carrier (Invitrogen, Carlsbad, CA) and 0.3 M sodium acetate. The pellet was re-suspended in 10 µl water to which was added 4 µl of 2X TANGO® buffer (MBI Fermentas, Inc.) and 1 µl of 10 U/µl *Sal* I (MBI Fermentas). Digestion was carried out a 37° C for 30 minutes. The reaction was terminated by incubation at 65° C for 15 minutes followed by ethanol precipitation in the presence of 0.3 M sodium acetate. The resulting pellet was resuspended in 10 µl water to which was added 4 µl of 2X TANGO® buffer and 1 µl of 10 U/µl *Bam* HI (MBI Fermentas, Inc.). Digestion was carried out at 37° C for 30 minutes. The reaction was terminated by extraction with phenol and chloroform followed by ethanol precipitation in the presence of 0.3 M sodium acetate.

The isolated cDNA inserts described above were ligated into *Xho* I and *Bam* HI digested and dephosphorylated *pZP9CEE* vector. The ligated DNA was transfected into MAXIMUM EFFICIENCY *DH10B*® competent cells (Life Technologies, Gaithersburg, MD). The resultant plasmid *pSLZchm1-8* encodes full-length native Zchm-1 polypeptide.

To facilitate detection and purification of the recombinant protein, a Glu tag was added to the C-terminal of Zchm-1. The native *Zchm-1* termination codon was removed by sited directed mutagenesis allowing the production of an inframe fusion of the Zchm-1 C-terminus to a Glu affinity tag encoded by *pZPCEE* vector sequences. A cDNA fragment encoding the mutagenized *Zchm-1* coding sequence was obtained by polymerase chain reaction from 30 ng of *Xho* I digested *pSLZchm103* template employing 5' primer SEQ ID NO:7 and 3' mutagenic primer *ZC18987* (SEQ ID NO:9). To

facilitate cloning into *pZP9CEE*, primers SEQ ID NO: 7 and SEQ ID NO:9 incorporate a *Sal* I and *Bam* HI restriction site, respectively. The conditions for polymerase chain reaction and cloning into *pZP9CEE* are identical to that described above for the construction of *pSLZChm1-8*. The resulting plasmid, *pSLZChm1CT-1*, encodes full-length Zchm-1 polypeptide with a C-terminal Glu tag.

Example 2

Northern Blot Analysis

A 683 bp *Eco*RI/*Bgl* II fragment purified from *pSLZchm1-3* was labeled with ^{32}P by random priming. The labeled *Zchm-1* probe was hybridized to multiple tissue Northern Blots (Clontech). The blots were washed at 50° C in 0.1X standard sodium citrate (SSC) buffer and exposed to X-ray film for 5 days with intensifying screens. Northern blot analysis shows a 1.4 kb *Zchm-1* transcript present in poly A mRNA samples purified from skeletal muscle and bone marrow.

Example 3

Chromosomal Localization

The *Zchm-1* locus is mapped to chromosome *11p15.4*. It was mapped to chromosome 11 using the "GeneBridge 4 Radiation Hybrid Panel" (Research Genetics, Inc. , Huntsville, AL). The GeneBridge 4 Radiation Hybrid Panel contains PCR amplifiable DNAs from each of 93 radiation hybrid clones, plus two control DNAs (the HFL donor and the A23 recipient). A publicly available WWW server (http://www-genome.wi.mit.edu/cgi-bin/contig/rhmapper.pl) allows mapping relative to the Whitehead Institute/MIT center for Genome Research's

radiation hybrid map of the human genome (the "WICGR" radiation hybrid map) which was constructed with the GeneBridge 4 Radiation Hybrid Panel.

For the mapping of *Zchm-1* with the "GeneBridge 4 RH Panel", 20 µl reactions were set up in a PCR amplifiable 96-well microtiter plate (Stratagene, La Jolla, CA) and used in a "RoboCycler Gradient 96" thermal cycler (Stratagene). Each of the 95 PCR reactions consisted of 2 µl 10X KlenTaq PCR reaction buffer (Clontech Laboratories, Inc., Palo Alto, CA), 1.6 µl dNTPs mix (2.5 mM each, PE Applied Biosystems, Foster City, CA), 1 µl sense primer ZC 18,314 5' CCG CGT CTG TGA ACC TTT 3' (SEQ ID NO:13), 1 µl antisense primer, ZC 18, 315, 5' GGG CCA CCC ACC AGT TAC 3' (SEQ ID NO:14), 2 µl REDILOAD® (Research Genetics, Inc.), 0.4 µl 50X Advantage KlenTaq Polymerase Mix (Clontech Laboratories, Inc.), 25 ng of DNA from an individual hybrid clone or control and water for a total volume of 20 µl. The reactions were overlaid with an equal amount of mineral oil and sealed. The PCR cycler conditions were as follows: an initial 1 cycle 5 minute denaturation at 95° C, 40 cycles of a 1 minute denaturation at 95° C, 1 minute annealing at 66° C and 1.5 minute extension at 72° C, followed by a final 1 cycle extension of 7 minutes at 72°C. The reactions were separated by electrophoresis on a 2% agarose gel (Life Technologies, Inc., Gaithersburg, MD).

The results showed that *Zchm-1* maps 17.08 cR_3000 from the framework marker *D11S922* on the chromosome 11 WICGR radiation hybrid map. Proximal and distal framework markers were *D11S922* and *D11S932*, respectively. The use of surrounding markers positions *Zchm-1* in the *11p15.4* region on the integrated LDB chromosome 11 map (The Genetic

Location Database, University of Southhampton, WWW server: http//cedar.genetics.soton.ac.uk/public_html/).

CLAIMS:

1. An isolated polynucleotide which encodes a polypeptide selected from the group consisting of SEQ ID NOs: 2, 3, 4, 5, 6, 11 and 12.

2. A polypeptide selected from the group consisting of SEQ ID NOs: 2, 3, 4, 5, 6, 11 and 12.

3. An antibody which specifically binds to a polypeptide selected from the group consisting of SEQ ID NOs: 2, 3, 4, 5, 6, 11 and 12.

SEQUENCE LISTING

<110> ZymoGenetics, Inc.

<120> Mammalian Chondromodulin-like Protein

<130> 98-74PC

<150> 09/191,986
<151> 1998-11-13

<160> 22

<170> FastSEQ for Windows Version 3.0

<210> 1
<211> 1178
<212> DNA
<213> Homo sapien

<220>
<221> CDS
<222> (58)...(1008)

<400> 1

```
agcagtggtc tctcagtcct ctcaaagcaa ggaaagagta ctgtgtgctg agagacc atg     60
                                                                 Met
                                                                  1

gca aag aat cct cca gag aat tgt gaa gac tgt cac att cta aat gca      108
Ala Lys Asn Pro Pro Glu Asn Cys Glu Asp Cys His Ile Leu Asn Ala
            5                  10                  15

gaa gct ttt aaa tcc aag aaa ata tgt aaa tca ctt aag att tgt gga      156
Glu Ala Phe Lys Ser Lys Lys Ile Cys Lys Ser Leu Lys Ile Cys Gly
        20                  25                  30

ctg gtg ttt ggt atc ctg gcc cta act cta att gtc ctg ttt tgg ggg      204
Leu Val Phe Gly Ile Leu Ala Leu Thr Leu Ile Val Leu Phe Trp Gly
    35                  40                  45

agc aag cac ttc tgg ccg gag gta ccc aaa aaa gcc tat gac atg gag      252
Ser Lys His Phe Trp Pro Glu Val Pro Lys Lys Ala Tyr Asp Met Glu
 50                  55                  60                  65
```

```
cac act ttc tac agc aat gga gag aag aag aag att tac atg gaa att      300
His Thr Phe Tyr Ser Asn Gly Glu Lys Lys Lys Ile Tyr Met Glu Ile
                70                  75                  80

gat cct gtg acc aga act gaa ata ttc aga agc gga aat ggc act gat      348
Asp Pro Val Thr Arg Thr Glu Ile Phe Arg Ser Gly Asn Gly Thr Asp
            85                  90                  95

gaa aca ttg gaa gtg cac gac ttt aaa aac gga tac act ggc atc tac      396
Glu Thr Leu Glu Val His Asp Phe Lys Asn Gly Tyr Thr Gly Ile Tyr
        100                 105                 110

ttc gtg ggt ctt caa aaa tgt ttt atc aaa act cag att aaa gtg att      444
Phe Val Gly Leu Gln Lys Cys Phe Ile Lys Thr Gln Ile Lys Val Ile
    115                 120                 125

cct gaa ttt tct gaa cca gaa gag gaa ata gat gag aat gaa gaa att      492
Pro Glu Phe Ser Glu Pro Glu Glu Glu Ile Asp Glu Asn Glu Glu Ile
130                 135                 140                 145

acc aca act ttc ttt gaa cag tca gtg att tgg gtc cca gca gaa aag      540
Thr Thr Thr Phe Phe Glu Gln Ser Val Ile Trp Val Pro Ala Glu Lys
                150                 155                 160

cct att gaa aac cga gat ttt ctt aaa aat tcc aaa att ctg gag att      588
Pro Ile Glu Asn Arg Asp Phe Leu Lys Asn Ser Lys Ile Leu Glu Ile
            165                 170                 175

tgt gat aac gtg acc atg tat tgg atc aat ccc act cta ata tca gtt      636
Cys Asp Asn Val Thr Met Tyr Trp Ile Asn Pro Thr Leu Ile Ser Val
        180                 185                 190

tct gag tta caa gac ttt gag gag gag gga gaa gat ctt cac ttt cct      684
Ser Glu Leu Gln Asp Phe Glu Glu Glu Gly Glu Asp Leu His Phe Pro
    195                 200                 205

gcc aac gaa aaa aaa ggg att gaa caa aat gaa cag tgg gtg gtc cct      732
Ala Asn Glu Lys Lys Gly Ile Glu Gln Asn Glu Gln Trp Val Val Pro
210                 215                 220                 225

caa gtg aaa gta gag aag acc cgt cac gcc aga caa gca agt gag gaa      780
Gln Val Lys Val Glu Lys Thr Arg His Ala Arg Gln Ala Ser Glu Glu
                230                 235                 240
```

```
gaa ctt cca ata aat gac tat act gaa aat gga ata gaa ttt gat ccc      828
Glu Leu Pro Ile Asn Asp Tyr Thr Glu Asn Gly Ile Glu Phe Asp Pro
            245                 250                 255

atg ctg gat gag aga ggt tat tgt tgt att tac tgc cgt cga ggc aac      876
Met Leu Asp Glu Arg Gly Tyr Cys Cys Ile Tyr Cys Arg Arg Gly Asn
        260                 265                 270

cgc tat tgc cgc cgc gtc tgt gaa cct tta cta ggc tac tac cca tat      924
Arg Tyr Cys Arg Arg Val Cys Glu Pro Leu Leu Gly Tyr Tyr Pro Tyr
    275                 280                 285

cca tac tgc tac caa gga gga cga gtc atc tgt cgt gtc atc atg cct      972
Pro Tyr Cys Tyr Gln Gly Gly Arg Val Ile Cys Arg Val Ile Met Pro
290                 295                 300                 305

tgt aac tgg tgg gtg gcc cgc atg ctg ggg agg gtc taataggagg          1018
Cys Asn Trp Trp Val Ala Arg Met Leu Gly Arg Val
                310                 315

tttgagctca aatgcttaaa ctgctggcaa catataataa atgcatgcta ttcaatgaat   1078
ttctgcctat gaggcatctg gcccctggta gccagctctc cagaattact tgtaggtaat   1138
tcctctcttc atgttctaat aaacttctac attatcacca                         1178
```

<210> 2
<211> 317
<212> PRT
<213> Homo sapien

<400> 2

```
Met Ala Lys Asn Pro Pro Glu Asn Cys Glu Asp Cys His Ile Leu Asn
 1               5                  10                  15
Ala Glu Ala Phe Lys Ser Lys Lys Ile Cys Lys Ser Leu Lys Ile Cys
            20                  25                  30
Gly Leu Val Phe Gly Ile Leu Ala Leu Thr Leu Ile Val Leu Phe Trp
        35                  40                  45
Gly Ser Lys His Phe Trp Pro Glu Val Pro Lys Lys Ala Tyr Asp Met
    50                  55                  60
Glu His Thr Phe Tyr Ser Asn Gly Glu Lys Lys Lys Ile Tyr Met Glu
65                  70                  75                  80
Ile Asp Pro Val Thr Arg Thr Glu Ile Phe Arg Ser Gly Asn Gly Thr
                85                  90                  95
Asp Glu Thr Leu Glu Val His Asp Phe Lys Asn Gly Tyr Thr Gly Ile
            100                 105                 110
Tyr Phe Val Gly Leu Gln Lys Cys Phe Ile Lys Thr Gln Ile Lys Val
```

```
        115                 120                 125
Ile Pro Glu Phe Ser Glu Pro Glu Glu Glu Ile Asp Glu Asn Glu Glu
    130                 135                 140
Ile Thr Thr Thr Phe Phe Glu Gln Ser Val Ile Trp Val Pro Ala Glu
145                 150                 155                 160
Lys Pro Ile Glu Asn Arg Asp Phe Leu Lys Asn Ser Lys Ile Leu Glu
                165                 170                 175
Ile Cys Asp Asn Val Thr Met Tyr Trp Ile Asn Pro Thr Leu Ile Ser
            180                 185                 190
Val Ser Glu Leu Gln Asp Phe Glu Glu Glu Gly Glu Asp Leu His Phe
        195                 200                 205
Pro Ala Asn Glu Lys Lys Gly Ile Glu Gln Asn Glu Gln Trp Val Val
    210                 215                 220
Pro Gln Val Lys Val Glu Lys Thr Arg His Ala Arg Gln Ala Ser Glu
225                 230                 235                 240
Glu Glu Leu Pro Ile Asn Asp Tyr Thr Glu Asn Gly Ile Glu Phe Asp
                245                 250                 255
Pro Met Leu Asp Glu Arg Gly Tyr Cys Cys Ile Tyr Cys Arg Arg Gly
            260                 265                 270
Asn Arg Tyr Cys Arg Arg Val Cys Glu Pro Leu Leu Gly Tyr Tyr Pro
        275                 280                 285
Tyr Pro Tyr Cys Tyr Gln Gly Gly Arg Val Ile Cys Arg Val Ile Met
    290                 295                 300
Pro Cys Asn Trp Trp Val Ala Arg Met Leu Gly Arg Val
305                 310                 315
```

<210> 3
<211> 15
<212> PRT
<213> Homo sapiens

<400> 3

```
Leu Val Phe Gly Ile Leu Ala Leu Thr Leu Ile Val Leu Phe Trp
 1               5                  10                  15
```

<210> 4
<211> 20
<212> PRT
<213> Homo sapiens

<400> 4

```
Ile Cys Gly Leu Val Phe Gly Ile Leu Ala Leu Thr Leu Ile Val Leu
 1               5                  10                  15
Phe Trp Gly Ser
            20
```

<210> 5
<211> 103
<212> PRT
<213> Homo sapiens

<400> 5
Gly Ile Glu Gln Asn Glu Gln Trp Val Val Pro Gln Val Lys Val Glu
1 5 10 15
Lys Thr Arg His Ala Arg Gln Ala Ser Glu Glu Glu Leu Pro Ile Asn
20 25 30
Asp Tyr Thr Glu Asn Gly Ile Glu Phe Asp Pro Met Leu Asp Glu Arg
35 40 45
Gly Tyr Cys Cys Ile Tyr Cys Arg Arg Gly Asn Arg Tyr Cys Arg Arg
50 55 60
Val Cys Glu Pro Leu Leu Gly Tyr Tyr Pro Tyr Pro Tyr Cys Tyr Gln
65 70 75 80
Gly Gly Arg Val Ile Cys Arg Val Ile Met Pro Cys Asn Trp Trp Val
85 90 95
Ala Arg Met Leu Gly Arg Val
100

<210> 6
<211> 63
<212> PRT
<213> Homo sapiens

<400> 6
Phe Asp Pro Met Leu Asp Glu Arg Gly Tyr Cys Cys Ile Tyr Cys Arg
1 5 10 15
Arg Gly Asn Arg Tyr Cys Arg Arg Val Cys Glu Pro Leu Leu Gly Tyr
20 25 30
Tyr Pro Tyr Pro Tyr Cys Tyr Gln Gly Gly Arg Val Ile Cys Arg Val
35 40 45
Ile Met Pro Cys Asn Trp Trp Val Ala Arg Met Leu Gly Arg Val
50 55 60

<210> 7
<211> 40
<212> DNA
<213> Homo sapiens

<400> 7
cacgtcgaca ccatggcaaa gaatcctcca gagaattgtg 40

<210> 8
<211> 40
<212> DNA
<213> Homo sapiens

<400> 8
acaggatccc tattagaccc tccccagcat gcgggccacc 40

<210> 9
<211> 36
<212> DNA
<213> Homo sapiens

<400> 9
acaggatccg accctcccca gcatgcgggc caccca 36

<210> 10
<211> 4
<212> PRT
<213> Homo sapiens

<400> 10
Arg Glu Arg Arg
1

<210> 11
<211> 110
<212> PRT
<213> Homo sapiens

<400> 11
Glu Gly Glu Asp Leu His Phe Pro Ala Asn Glu Lys Lys Gly Ile Glu
1 5 10 15
Gln Asn Glu Gln Trp Val Val Pro Gln Val Lys Val Glu Lys Thr Arg
20 25 30
His Ala Arg Gln Ala Ser Glu Glu Glu Leu Pro Ile Asn Asp Tyr Thr
35 40 45
Glu Asn Gly Ile Glu Phe Asp Pro Met Leu Asp Glu Arg Gly Tyr Cys
50 55 60
Cys Ile Tyr Cys Arg Arg Gly Asn Arg Tyr Cys Arg Arg Val Cys Glu
65 70 75 80
Pro Leu Leu Gly Tyr Tyr Pro Tyr Pro Tyr Cys Tyr Gln Gly Gly Arg
85 90 95
Val Ile Cys Arg Val Ile Met Pro Cys Asn Trp Trp Val Ala
100 105 110

<210> 12
<211> 264
<212> PRT
<213> Homo sapiens

<400> 12

```
Trp Gly Ser Lys His Phe Trp Pro Glu Val Pro Lys Lys Ala Tyr Asp
 1               5                  10                  15
Met Glu His Thr Phe Tyr Ser Asn Gly Glu Lys Lys Lys Ile Tyr Met
            20                  25                  30
Glu Ile Asp Pro Val Thr Arg Thr Glu Ile Phe Arg Ser Gly Asn Gly
        35                  40                  45
Thr Asp Glu Thr Leu Glu Val His Asp Phe Lys Asn Gly Tyr Thr Gly
    50                  55                  60
Ile Tyr Phe Val Gly Leu Gln Lys Cys Phe Ile Lys Thr Gln Ile Lys
65                  70                  75                  80
Val Ile Pro Glu Phe Ser Glu Pro Glu Glu Glu Ile Asp Glu Asn Glu
                85                  90                  95
Glu Ile Thr Thr Thr Phe Phe Glu Gln Ser Val Ile Trp Val Pro Ala
            100                 105                 110
Glu Lys Pro Ile Glu Asn Arg Asp Phe Leu Lys Asn Ser Lys Ile Leu
        115                 120                 125
Glu Ile Cys Asp Asn Val Thr Met Tyr Trp Ile Asn Pro Thr Leu Ile
    130                 135                 140
Ser Val Ser Glu Leu Gln Asp Phe Glu Glu Glu Gly Glu Asp Leu His
145                 150                 155                 160
Phe Pro Ala Asn Glu Lys Lys Gly Ile Glu Gln Asn Glu Gln Trp Val
                165                 170                 175
Val Pro Gln Val Lys Val Glu Lys Thr Arg His Ala Arg Gln Ala Ser
            180                 185                 190
Glu Glu Glu Leu Pro Ile Asn Asp Tyr Thr Glu Asn Gly Ile Glu Phe
        195                 200                 205
Asp Pro Met Leu Asp Glu Arg Gly Tyr Cys Cys Ile Tyr Cys Arg Arg
    210                 215                 220
Gly Asn Arg Tyr Cys Arg Arg Val Cys Glu Pro Leu Leu Gly Tyr Tyr
225                 230                 235                 240
Pro Tyr Pro Tyr Cys Tyr Gln Gly Gly Arg Val Ile Cys Arg Val Ile
                245                 250                 255
Met Pro Cys Asn Trp Trp Val Ala
            260
```

<210> 13
<211> 18
<212> DNA

<213> Homo sapiens

<400> 13
ccgcgtctgt gaaccttt 18

<210> 14
<211> 18
<212> DNA
<213> Homo sapiens

<400> 14
gggccaccca ccagttac 18

<210> 15
<211> 248
<212> DNA
<213> Homo sapiens

<400> 15
cagcagtngt ctctcagtcc tctcaaagca aggaaagagt actgtgtgct gagagaccat 60
ggcaaagaat cctccagaga attgtgaaga ctgtcacatt ctaaatgcag aagcttttaa 120
atccaagaaa atatgtaaat cacttaagat ttgtggactg gtgtttggta tcctggccct 180
aactctaatt gtcctgtttt gggggagcaa gcacttctgg ccggaggtac ccaaaaaagc 240
ctatgaca 248

<210> 16
<211> 28
<212> PRT
<213> Homo sapiens

<400> 16
Lys Asn Pro Pro Glu Asn Cys Glu Asp Cys His Ile Leu Asn Ala Glu
1 5 10 15
Ala Phe Lys Ser Lys Lys Ile Cys Lys Ser Leu Lys
20 25

<210> 17
<211> 61
<212> PRT
<213> Homo sapiens

<400> 17
Ser Lys His Phe Trp Pro Glu Val Pro Lys Lys Ala Tyr Asp Met Glu
1 5 10 15
His Thr Phe Tyr Ser Asn Gly Glu Lys Lys Lys Ile Tyr Met Glu Ile

```
            20                  25                  30
Asp Pro Val Thr Arg Thr Glu Ile Phe Arg Ser Gly Asn Gly Thr Asp
        35                  40                  45
Glu Thr Leu Glu Val His Asp Phe Lys Asn Gly Tyr Thr
    50                  55                  60
```

<210> 18
<211> 50
<212> PRT
<213> Homo sapiens

<400> 18

```
Ser Glu Pro Glu Glu Glu Ile Asp Glu Asn Glu Glu Ile Thr Thr Thr
 1               5                  10                  15
Phe Phe Glu Gln Ser Val Ile Trp Val Pro Ala Glu Lys Pro Ile Glu
            20                  25                  30
Asn Arg Asp Phe Leu Lys Asn Ser Lys Ile Leu Glu Ile Cys Asp Asn
        35                  40                  45
Val Thr
    50
```

<210> 19
<211> 87
<212> PRT
<213> Homo sapiens

<400> 19

```
Glu Leu Gln Asp Phe Glu Glu Glu Gly Glu Asp Leu His Phe Pro Ala
 1               5                  10                  15
Asn Glu Lys Lys Gly Ile Glu Gln Asn Glu Gln Trp Val Val Pro Gln
            20                  25                  30
Val Lys Val Glu Lys Thr Arg His Ala Arg Gln Ala Ser Glu Glu Glu
        35                  40                  45
Leu Pro Ile Asn Asp Tyr Thr Glu Asn Gly Ile Glu Phe Asp Pro Met
    50                  55                  60
Leu Asp Glu Arg Gly Tyr Cys Cys Ile Tyr Cys Arg Arg Gly Asn Arg
65                  70                  75                  80
Tyr Cys Arg Arg Val Cys Glu
                85
```

<210> 20
<211> 54
<212> PRT
<213> Homo sapiens

<400> 20
Lys Val Glu Lys Thr Arg His Ala Arg Gln Ala Ser Glu Glu Glu Leu
1 5 10 15
Pro Ile Asn Asp Tyr Thr Glu Asn Gly Ile Glu Phe Asp Pro Met Leu
20 25 30
Asp Glu Arg Gly Tyr Cys Cys Ile Tyr Cys Arg Arg Gly Asn Arg Tyr
35 40 45
Cys Arg Arg Val Cys Glu
50

<210> 21
<211> 57
<212> PRT
<213> Homo sapiens

<400> 21
Glu Leu Gln Asp Phe Glu Glu Glu Gly Glu Asp Leu His Phe Pro Ala
1 5 10 15
Asn Glu Lys Lys Gly Ile Glu Gln Asn Glu Gln Trp Val Val Pro Gln
20 25 30
Val Lys Val Glu Lys Thr Arg His Ala Arg Gln Ala Ser Glu Glu Glu
35 40 45
Leu Pro Ile Asn Asp Tyr Thr Glu Asn
50 55

<210> 22
<211> 103
<212> PRT
<213> Homo sapiens

<400> 22
Glu Lys Pro Ile Glu Asn Arg Asp Phe Leu Lys Asn Ser Lys Ile Leu
1 5 10 15
Glu Ile Cys Asp Asn Val Thr Met Tyr Trp Ile Asn Pro Thr Leu Ile
20 25 30
Ser Val Ser Glu Leu Gln Asp Phe Glu Glu Glu Gly Glu Asp Leu His
35 40 45
Phe Pro Ala Asn Glu Lys Lys Gly Ile Glu Gln Asn Glu Gln Trp Val
50 55 60
Val Pro Gln Val Lys Val Glu Lys Thr Arg His Ala Arg Gln Ala Ser
65 70 75 80
Glu Glu Glu Leu Pro Ile Asn Asp Tyr Thr Glu Asn Gly Ile Glu Phe
85 90 95
Asp Pro Met Leu Asp Glu Arg
100

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 99/26909

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9953051 A	21-10-1999	AU 3050199 A	01-11-1999

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 99/26909

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/12 C07K14/51 C07K16/22

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	WO 99 53051 A (DUCLERT AYMERIC ;GENSET (FR); DUMAS MILNE EDWARDS JEAN BAPTI (FR);) 21 October 1999 (1999-10-21) SEQ.ID. 408 and claims	1-3
X	HILLIER ET AL.: "WashU-Merck EST Project." EMBL DATABASE ENTRY HS1154447; ACCESSION NUMBER AA236166,6 March 1997 (1997-03-06), XP002133435 abstract	1
A	STRAUSBERG R.: "National Cancer Institute, Cancer Genome Anatomy Project." EMBL DATABASE ENTRY AI123839; ACCESSION NUMBER AI123839, 8 September 1998 (1998-09-08), XP002133436 abstract	1-3

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
17 March 2000	07.04.00

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Mandl, B

1

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	LIEW C. C. ET AL.: "A catalogue of genes in the cardiovascular system as identified by expressed sequence tags." EMBL DATABASE ENTRY HS1807; ACCESSION NUMBER T12180, 18 December 1994 (1994-12-18), XP002133437 the whole document ---	1-3
A	ADAMS M. D. ET AL.: "Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence." EMBL DATABASE ENTRY HSZZ02372; ACCESSION NUMBER AA297231, 18 April 1997 (1997-04-18), XP002133438 abstract -----	1-3